

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 12:55:11 ; Search time 36 Seconds
(without alignments)
762.490 Million cell updates/sec

Title: US-09-646-778A-288
Perfect score: 1063
Sequence: 1 RUSCATLGSQPHPSRLT.....DTALLDNMKALKLKTTEL 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	A_Geneseq_101002.*
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1063	100.0	206	20	AAV73837 Human prostate tum
2	1002	94.3	194	20	AAV76590 Human ovarian tumo
3	1002	94.3	202	22	AAV75621 Human colon cancer
4	899	84.6	175	19	AAV77365 zsig10 polypeptide
5	899	84.6	175	19	AAV37844 Human XAG growth f
6	899	84.6	175	19	AAV37844 Human protein comp
7	899	84.6	175	20	AAV59675 Secreted protein 1
8	899	84.6	175	21	AAV24070 Human PRO1030 prot
9	899	84.6	175	22	AAU08604 Breast cancer cell
10	899	84.6	175	22	AAV72203 Human huXAG-1/CCSG

11	899	84.6	175	23	ABB78994 Human XAG-2 protei
12	899	84.6	175	23	ABG34043 Human pro peptide
13	899	84.6	175	23	AAU74939 Human clone DNA594
14	830	78.1	175	23	AAU83628 Human PRO protein,
15	656	61.7	132	21	AAU00109 Human secreted pro
16	605	56.9	131	20	AAV59718 Secreted protein 7
17	598	56.3	116	20	AAV12312 Human 5' EST relat
18	593	55.8	115	21	AAV64672 Human XAG growth f
19	555	52.2	166	19	AAW37846 Breast cancer prot
20	555	52.2	166	21	AAU00194 Breast cancer-asso
21	555	52.2	166	22	AAU25728 Human huXAG-3/CCSG
22	555	52.2	166	22	AAU72205 Amino acid sequenc
23	555	52.2	166	22	AAU31192 Human BCMP 11 poly
24	555	52.2	166	22	AAU07647 BCMP-11 protein.
25	555	52.2	166	23	AAU23228 Human angiogenesis
26	555	52.2	166	23	ABB95520 Human PRO1272 prot
27	555	52.2	166	23	ABB84914 Human PRO protein,
28	555	52.2	166	23	AAU83649 Colon tumour relat
29	555	52.2	168	22	AAU24502 Human ovarian anti
30	555	52.2	184	23	ABP41417 Human ovarian tumo
31	388	36.5	89	20	AAV76568 Human colon cancer
32	361	34.0	180	22	AAU73737 Human 5' EST secre
33	285	26.8	56	20	AAV11881 Human 5' EST secre
34	222	20.9	56	20	AAV11939 Human XAG growth f
35	222	20.9	172	19	AAW37845 Clone HP01766 of a
36	222	20.9	172	20	AAV03236 Amino acid sequenc
37	222	20.9	172	20	AAV03237 Xenopus sp embryo
38	222	20.9	172	20	AAV03230 Human PRO1376 prot
39	222	20.9	172	21	AAU10282 Human membrane or
40	222	20.9	172	21	AAU24432 Human huXAG-2/CCSG
41	222	20.9	172	22	AAU88366 Human huXAG-2/CCSG
42	222	20.9	172	22	AAU72204 Amino acid sequenc
43	222	20.9	172	23	ABV55536 Human angiogenesis
44	218.5	20.6	172	19	AAU80812 Amino acid sequenc
45	214	20.1	172	23	ABB84930 Human PRO1376 prot

ALIGNMENTS

RESULT 1	AAV73837
ID	AAV73837 standard; Protein; 206 AA.
XX	AAV73837;
XX	14-MAR-2000 (first entry)
DT	Human prostate tumor EST-fragment derived protein #24.
XX	Pancreas; tumor; EST; expressed sequence tag; human; cytostatic; treatment.
XX	Homo sapiens.
OS	DE19820190-AL.
PN	04-NOV-1999.
XX	28-APR-1998; 98DE-1020190.
PF	28-APR-1998; 98DE-1020190.
XX	(META-) METAGEN GES GENOMFORSCHUNG MBH.
XX	Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
PI	WPI; 1999-621386/54.
XX	N-PSDB; AAZ52865.
DR	New human nucleic acid sequences from pancreatic tumors, and related proteins
XX	

CC	AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC	cancer-associated nucleic acid molecules (N) and proteins (P), where
CC	the proteins are collectively known as colon cancer antigens. The colon
CC	cancer antigens have cytostatic activity and can be used in gene
CC	therapy and vaccine production. N and P may be used in the prevention,
CC	diagnosis and treatment of diseases associated with inappropriate P
CC	expression. For example, N and P may be used to treat disorders
CC	associated with decreased expression by rectifying mutations or deletions
CC	in a patient's genome that affect the activity of P by expressing
CC	inactive proteins or to supplement the patients own production of P.
CC	Additionally, N may be used to produce the colon cancer-associated Ps,
CC	by inserting the nucleic acids into a host cell and culturing the cell
CC	to express the proteins. N and P can be used in the prevention, diagnosis
CC	and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC	and AAB77789 represent sequences used in the exemplification of the
CC	present invention.
CC	N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC	missing at time of publication, meaning no sequences are present for
CC	SEQ ID NO:1027 to 1052, 7921 and 7922.
XX	
SQ	Sequence 202 AA;
Query Match 94.3%; Score 1002; DB 22; Length 202;	
Best Local Similarity 100.0%; Pred. No. 1.3e-101;	
Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	13 PHSRRLTGGRWKRVRAMEKIPVSAPLLLVLSYTLARDTTVPKAKKDKDSRPKLP 72
DB	9 PHSRRLTGGRWKRVRAMEKIPVSAPLLLVLSYTLARDTTVPKAKKDKDSRPKLP 68
QY	73 QTLSRGWDQLIWQTVEEALYKSKTSNKPLMIHHLDCEPHSQALKKVFNAENKEIQKLA 132
DB	69 QTLSRGWDQLIWQTVEEALYKSKTSNKPLMIHHLDCEPHSQALKKVFNAENKEIQKLA 128
QY	133 EQVLLNLVYETTDKHLSPDQGYVPRIMFVDPSTLVRADITGRYSNRLYAYEPADTALL 192
DB	129 EQVLLNLVYETTDKHLSPDQGYVPRIMFVDPSTLVRADITGRYSNRLYAYEPADTALL 188
QY	193 DNKKALKLKTTEL 206
DB	189 DNKKALKLKTTEL 202
RESULT 4	
AAW77365	
ID	AAW77365 standard; Protein; 175 AA.
XX	
AC	AAW77365;
XX	
DT	21-DEC-1998 (first entry)
XX	
DE	zsig10 polypeptide.
XX	
KW	Human; mucous-mediated function; adhesion; tumour metastasis;
KW	bacterial colonisation; microbial infection; AIDS; cystic fibrosis;
KW	chronic obstructive pulmonary disease; asthma; Crohn's disease;
KW	sinonasal inflammatory disease; inflammatory bowel disease; bronchitis.
XX	
OS	Homo sapiens.
XX	
PN	WO9841627-A1.
XX	
PD	24-SEP-1998.
XX	
PF	18-MAR-1998; 98WO-US05251.
XX	
PR	19-MAR-1997; 97US-0039631.
XX	
PA	(ZYMO) ZYMOGENETICS INC.
XX	
PT	Sheppard PO;
XX	
DR	WPI; 1998-531566/45.
N-PSDB; AAV59320.	
XX	
PT	New isolated mucous-associated polypeptide, zsig10 - used to develop
PT	products for treating e.g. tumour metastasis, microbial infections,
PT	cystic fibrosis, asthma, bronchitis or inflammatory bowel disease
XX	
PS	Claim 1; Page 82; 109pp; English.
XX	
CC	The human polypeptide zsig10 is involved in mucous-mediated functions
CC	such as adhesion. The products of the invention can be used in the study
CC	and treatment of e.g. tumour metastasis, bacterial colonisation,
CC	susceptibility to and persistence of infection, microbial infections,
CC	AIDS, cystic fibrosis, chronic obstructive pulmonary disease, asthma,
CC	sinonasal inflammatory disease, inflammatory bowel disease, bronchitis,
CC	or Crohn's disease. The products can also be used for detection,
CC	diagnosis and drug screening.
XX	
SQ	Sequence 175 AA;
Query Match 84.6%; Score 899; DB 19; Length 175;	
Best Local Similarity 100.0%; Pred. No. 2.2e-90;	
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	32 MEKIPVSAFLLLVALSYTLARDTTVPKAKKDKDSRPKLPQTLSRGWDQLIWQTVEE 91
DB	1 MEKIPVSAFLLLVALSYTLARDTTVPKAKKDKDSRPKLPQTLSRGWDQLIWQTVEE 60
QY	92 ALYKSKTSNKPLMIHHLDCEPHSQALKKVFNAENKEIQKLAEQVLLNLVYETTDKHLSP 151
DB	61 ALYKSKTSNKPLMIHHLDCEPHSQALKKVFNAENKEIQKLAEQVLLNLVYETTDKHLSP 120
QY	152 DGQYVPRIMFVDPSTLVRADITGRYSNRLYAYEPADTALLDNMKALKLKTTEL 206
DB	121 DGQYVPRIMFVDPSTLVRADITGRYSNRLYAYEPADTALLDNMKALKLKTTEL 175
RESULT 5	
AAW37844	
ID	AAW37844 standard; Protein; 175 AA.
XX	
AC	AAW37844;
XX	
DT	28-AUG-1998 (first entry)
XX	
DE	Human XAG growth factor huXAG-1.
XX	
KW	HuXAG-1; XAG; growth factor; colon cancer; tumour marker;
KW	breast disease; liver disease; lung disease; emphysema;
KW	wound healing; diagnosis; therapy; human.
XX	
OS	Homo sapiens.
XX	
FT	Key Location/Qualifiers
FT	Peptide 1..20
FT	/label= sig_peptide
FT	Protein 21..175
FT	/label= Mat_protein
FT	/note= "Claim 11"
FT	Peptide 20..43
FT	/label= Epitope
FT	/note= "Claim 12"
FT	Peptide 44..52
FT	/label= Epitope
FT	/note= "Claim 12"
FT	Peptide 61..72
FT	/label= Epitope
FT	/note= "Claim 12"
FT	Peptide 90..103
FT	/label= Epitope
FT	/note= "Claim 12"
FT	Peptide 113..125
FT	/label= Epitope
FT	/note= "Claim 12"

FT Peptide 138...150
 FT /label= Epitope
 FT /note= "Claim 12"
 PN WO9807749-A1.
 XX 26-FEB-1998.
 XX 22-AUG-1997; 97WO-US14139.
 XX 23-AUG-1996; 96WO-US13766.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Dillion PJ, Ebner R, Endress GA, Yu G;
 XX WPI: 1998-169093/15.
 XX N-PSDB; AAV19155.
 XX New isolated human XAG growth factor(s) - used to develop products
 XX for treating e.g. liver, lung or breast diseases or
 XX hyperproliferative disorders, e.g. cancer.
 XX Claim 1; Fig 1; 141pp; English.
 XX This polypeptide comprises huxAG-1, a member of a novel family of
 CC human growth factors also including huxAG-2 (see AAW37845) and
 CC huxAG-3 (see AAW37846). These proteins share homology with the
 CC XAG protein of *Xenopus laevis*, which is involved in embryogenesis
 CC and is expressed in adult tissue. huxAG-1 is specifically found
 CC in cancerous colon cells and may therefore be a growth factor for
 CC colon cancer. huxAG-1 cDNA (see AAV19155) was isolated from a cDNA
 CC library derived from human colon cancer tissue. Vectors, host
 CC cells, antibodies, and screening methods for identifying agonists
 CC and antagonists of huxAG-1 are provided. huxAG polypeptides are
 CC growth factors and can be used to stimulate proliferation of cells.
 CC They can be used to stimulate the proliferation and differentiation
 CC of hepatocytes to alleviate or treat liver diseases and pathologies
 CC such as fulminant liver failure caused by cirrhosis, liver damage
 CC caused by viral hepatitis and toxic substances. They can also be
 CC used to stimulate or promote liver regeneration, e.g. after
 CC surgery. They can also be used to prevent and heal damage to the
 CC lungs caused by various pathological states. They can be used to
 CC stimulate proliferation and differentiation and promote the repair
 CC of alveoli and bronchiolar epithelium to prevent, attenuate, or
 CC treat acute or chronic lung damage, e.g. emphysema, which results
 CC in the progressive loss of alveoli, and inhalation injuries, e.g.
 CC resulting from smoke inhalation and burns, that cause necrosis of
 CC the bronchiolar epithelium and alveoli. They can also be used to
 CC stimulate the proliferation and differentiation of breast tissue
 CC and could therefore be used to promote healing of breast tissue
 CC injury due to surgery, trauma or cancer. Antagonists can be used
 CC to treat hyperproliferative disorders, including cancer, in
 CC particular hepatocellular carcinoma, osteoclastoma, breast cancer,
 CC or colon cancer. The products can also be used for detection and
 CC diagnosis.

Query Match 84.6%; Score 899; DB 19; Length 175;
 Best Local Similarity 100.0%; Pred. No. 2.2e-90;
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 MEKIPVSAPLLLVSYTLARDTTVPKAKDKTDSRPKLPQTLSRGWGDLIWTQYEE 91
 DB 1 MEKIPVSAPLLLVSYTLARDTTVPKAKDKTDSRPKLPQTLSRGWGDLIWTQYEE 60

OY 92 ALYKSTSNKPLMIHHLDECPSHQALKKVFENKETQKLAEOVLLNLVYETTDKHLSP 151
 DB 61 ALYKSTSNKPLMIHHLDECPSHQALKKVFENKETQKLAEOVLLNLVYETTDKHLSP 120

OY 152 DGQYVPRIMFVDPSTVRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 206
 DB 152 DGQYVPRIMFVDPSTVRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 175

RESULT 7
 AAY59675

Db 121 DGQYVPRIMFVDPSTVRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 175

RESULT 6
 AAW37872
 ID AAW37872 standard; Protein; 175 AA.
 XX AAW37872;
 AC AAW37872;
 XX 10-AUG-1998 (first entry)
 DT 10-AUG-1998 (first entry)
 XX Human protein comprising secretory signal amino acid sequence 9.
 DE Human protein; secretory signal; nutritional source; cytokine;
 XX immunity/haematopoiesis; activin; inhibitor; tumour; chemotactic;
 KW chemokinetic; thrombolytic; anti-inflammatory; inhibition;
 KW stomach cancer cell.
 XX Homo sapiens.
 OS Homo sapiens.
 XX W09811217-A2.
 PN 19-MAR-1998.
 XX 12-SEP-1997; 97WO-JP03239.
 XX 13-SEP-1996; 96JP-0243060.
 XX (PROT-) PROTEGENE INC.
 PA (SAGA) SAGAMI CHEM RES CENTRE.
 XX Kato S, Kobayashi M, Sekine S, Yamaguchi T;
 PI WPI: 1998-207380/18.
 XX N-PSDB; AAV29047, AAV29048.
 DR Human proteins with secretory signal sequences - used to treat
 PT immune deficiencies, infections, tumours, and haematopoietic
 PT disorders, etc.
 XX Claim 1; Pages 79; 131pp; English.
 PS This is the amino acid sequence of a novel human protein comprising
 CC a secretory signal isolated from stomach cancer cells. Its proteins
 CC can be used as nutritional sources or supplements. The proteins may
 CC also have cytokine functions, immune modulating functions,
 CC haematopoiesis regulating activity, activin/inhibin regulating
 CC activity, chemotactic/chemokinetic activity, haemostatic and
 CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, tumour inhibition activity.

Query Match 84.6%; Score 899; DB 19; Length 175;
 Best Local Similarity 100.0%; Pred. No. 2.2e-90;
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 MEKIPVSAPLLLVSYTLARDTTVPKAKDKTDSRPKLPQTLSRGWGDLIWTQYEE 91
 DB 1 MEKIPVSAPLLLVSYTLARDTTVPKAKDKTDSRPKLPQTLSRGWGDLIWTQYEE 60

OY 92 ALYKSTSNKPLMIHHLDECPSHQALKKVFENKETQKLAEOVLLNLVYETTDKHLSP 151
 DB 61 ALYKSTSNKPLMIHHLDECPSHQALKKVFENKETQKLAEOVLLNLVYETTDKHLSP 120

OY 152 DGQYVPRIMFVDPSTVRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 206
 DB 152 DGQYVPRIMFVDPSTVRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 175

RESULT 7
 AAY59675

ID AAY59675 standard; Protein; 175 AA.
 XX AAY59675;
 AC AAY59675;
 XX 18-JAN-2000 (first entry)
 DT 18-JAN-2000 (first entry)
 DE Secreted protein 108-008-5-0-A6-FL.
 XX
 KW Secreted protein; fingerprint identification technique;
 KW chromosome mapping; human; hereditary disease; diagnosis; cancer;
 KW hyperlipidaemia; cardiovascular; neurodegenerative disorder; therapy;
 KW autoimmune disease; rheumatic disease; embryogenic disorder; myopathy;
 KW renal injury; amino aciduria; hypoglycaemia; male rat infertility;
 KW hypertension.
 XX Homo sapiens.
 OS Homo sapiens.
 XX WO9940189-A2.
 PN WO9940189-A2.
 XX 12-AUG-1999.
 PD 12-AUG-1999.
 XX 09-FEB-1999; 99WO-IB00282.
 PF 09-FEB-1999;
 XX 09-FEB-1998; 98US-0074121.
 PR 13-APR-1998; 98US-0081563.
 PR 10-AUG-1998; 98US-0096116.
 PR 04-SEP-1998; 98US-0099273.
 XX (GEST) GENSET.
 PA Bougueleret L, Duclert A, Dumas Milne Edwards J;
 XX WPI; 1999-600966/51.
 XX N-PSDB; AA240803.
 DR Extended cDNAs useful for expressing secreted proteins and to obtain
 DR specific antibodies -
 PT Claim 10; Page 199; 244pp; English.
 XX This sequence represents a human secreted protein of the invention.
 CC The extended cDNAs (or genomic DNAs obtainable from them) may be used to
 CC prepare PCR primers and probes. These are useful for forensic matching or
 CC positive identification by DNA sequencing. They may also be used in
 CC alternative fingerprint identification techniques. Antibodies against the
 CC proteins encoded by the extended cDNAs are useful in identification of
 CC tissue types or cell species, as well as identifying tissue specific
 CC soluble proteins. The sequences can be used for chromosome mapping and
 CC identification of genes associated with hereditary diseases or drug
 CC response. Signal sequences from the cDNAs can be used in construction of
 CC secretion vectors. Other sequences derived from the extended cDNAs can be
 CC used to clone upstream genomic DNA sequences including promoters. This is
 CC in turn useful for identifying proteins that interact with promoter
 CC sequences. Some of the proteins may be useful in diagnosing and treating
 CC several disorders including, but not limited to: cancer, hyperlipidaemia,
 CC cardiovascular and neurodegenerative disorders, autoimmune diseases, and
 CC rheumatic diseases, embryogenic disorders, hypertension, renal injury,
 CC amino acidurias, hypoglycaemia, male rat infertility and myopathies.
 XX Sequence 175 AA:
 SQ
 Query Match 84.6%; Score 899; DB 20; Length 175;
 Best Local Similarity 100.0%; Pred. No. 2.2e-90;
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 32 MEKIPVSALLLVLSYTLARDTTVKPGAKDQKDSRPKLQPTLSRGWGDLWTQTYEE 91
 1 MEKIPVSALLLVLSYTLARDTTVKPGAKDQKDSRPKLQPTLSRGWGDLWTQTYEE 60
 92 ALYKSKTSNKNPLMITHLDCPHSQALKKVFENKEIQKLAEOFVLLNLVYETDKHLSP 151
 61 ALYKSKTSNKNPLMITHLDCPHSQALKKVFENKEIQKLAEOFVLLNLVYETDKHLSP 120

QY 152 DQYVPRIMFVDPSTVRADITGRYSNRLIYAYEPADTALLDNMKKALKLLKTEL 206
 DB 121 DQYVPRIMFVDPSTVRADITGRYSNRLIYAYEPADTALLDNMKKALKLLKTEL 175
 RESULT 8
 AAB24070
 ID AAB24070 standard; Protein; 175 AA.
 XX AAB24070;
 AC AAB24070;
 XX 29-JAN-2001 (first entry)
 DT 29-JAN-2001 (first entry)
 DE Human PRO1030 protein sequence SEQ ID NO:40.
 XX Human; tumour; diagnosis; neoplastic disease; neoplastic cell growth;
 KW proliferation; tumorigenesis; identification; cancer; cytostatic;
 KW neotropic; neuroprotective; antiinflammatory; immunosuppressive;
 KW immunostimulant; antiangiogenic; leukaemia; lymphoid malignancy;
 KW neuronal disorder; glial disorder; astrocytal disorder; angiogenic;
 KW hypothalamic disorder; glandular disorder; macrophagal disorder;
 KW epithelial disorder; stromal disorder; blastocoealic disorder;
 KW inflammatory disorder; immunologic disorder.
 XX Homo sapiens.
 OS Homo sapiens.
 XX WO200053755-A2.
 PN WO200053755-A2.
 XX 14-SEP-2000.
 PD 14-SEP-2000.
 XX 06-JAN-2000; 2000WO-US00376.
 PF 06-JAN-2000;
 XX 08-MAR-1999; 99WO-US05028.
 PR 02-JUN-1999; 99WO-US12252.
 PR 23-JUN-1999; 99US-0141037.
 PR 07-JUL-1999; 99US-0143048.
 PR 26-JUL-1999; 99US-0145698.
 PR 30-NOV-1999; 99WO-US28313.
 PR 20-DEC-1999; 99WO-US30911.
 PR 05-JAN-2000; 2000WO-US00219.
 XX (GETH) GENENTECH INC.
 XX Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hillan KJ, Roy MA;
 PI Watanabe CK, Wood WI;
 XX WPI; 2000-572270/53.
 DR N-PSDB; AAC58380.
 XX Thirty PRO polynucleotides encoding PRO polypeptides, useful in the
 PT treatment, diagnosis and prevention of cancer -
 XX Claim 61; Fig 28; 286pp; English.
 XX The present invention describes an isolated antibody that binds to
 CC one of the human PRO proteins designated PRO212, PRO290, PRO341, PRO535,
 CC PRO619, PRO717, PRO809, PRO830, PRO848, PRO943, PRO1005, PRO1009,
 CC PRO1025, PRO1030, PRO1097, PRO1107, PRO1111, PRO1153, PRO1182, PRO1184,
 CC PRO1187, PRO1281, PRO23, PRO39, PRO834, PRO1317, PRO1710, PRO2094,
 CC PRO2145 OR PRO2198. PRO antagonists can be used to inhibit tumour cell
 CC growth. The PRO polypeptides and nucleotides are useful in the
 CC treatment, diagnosis and prevention of cancer. The antibodies and other
 CC anti-tumour compounds may be used to treat various conditions, including
 CC those characterised by overexpression and/or activation of the amplified
 CC PRO genes. Exemplary conditions or disorders to be treated with such
 CC antibodies and other compounds include benign or malignant tumours
 CC (e.g., renal, liver, kidney, bladder, breast, gastric, ovarian,
 CC colorectal, prostate, pancreatic, lung, vulva, thyroid, hepatic,
 CC carcinomas, sarcomas, glioblastomas, and various head and neck tumours),
 CC leukaemias and lymphoid malignancies, other disorders such as neuronal,
 CC glial, astrocytal, hypothalamic and other glandular, macrophagal,
 CC epithelial, stromal and blastocoealic disorders, and inflammatory,
 CC angiogenic and immunologic disorders. AAC58242 to AAC58366 represent PCR

The invention describes the novel use of a protein found in breast cancer cell membranes (BCMP 7) for diagnosing, preventing and treating breast cancers. The peptide has cytostatic action and potential uses in gene therapy and in vaccines. The polypeptide, antisense nucleic acids, or fusion proteins comprising and Green Fluorescent protein or the DsRed, Fluorescent-protein, antibodies specific for and/or nucleic acid are used for the prevention and/or treatment of breast cancer antibodies

CC for promoting healing of breast tissue injury due to surgery, trauma or
CC cancer. HuXAG-1 and the identified antagonist are useful for treating
CC cancer, in particular colon cancer. Detecting altered levels of huXAG-1
CC and its polynucleotides are useful for diagnosing or detecting cancer in
CC mammals. The gene encoding huXAG-1 is useful for monitoring human
CC colorectal carcinoma. huXAG-1 nucleic acid molecules are also useful for
CC chromosome identification. The present sequence represents the huXAG-1
CC protein.
XX
XX

SQ Sequence 175 AA;
Query Match 84.6%; Score 899; DB 22; Length 175;
Best Local Similarity 100.0%; Pred. No. 2.2e-90;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 32 MEKIPVSFAFLLLVALSYTLARDTTVPKAKKDTKDSRPKLPQTLSRGWGDLWTQYEE 91
DB 1 MEKIPVSFAFLLLVALSYTLARDTTVPKAKKDTKDSRPKLPQTLSRGWGDLWTQYEE 60
QY 92 ALYKSTSNKPLMIHHLDECPSHQAALKKVFPAENKEIOKLAEOQVLLNLVYETTDKHLSP 151
DB 61 ALYKSTSNKPLMIHHLDECPSHQAALKKVFPAENKEIOKLAEOQVLLNLVYETTDKHLSP 120
QY 152 DGOYVPRIMFVDPSTLRADITGRYSNRLYAYEPADTALLDDNMKKALKLKTTEL 206
DB 121 DGOYVPRIMFVDPSTLRADITGRYSNRLYAYEPADTALLDDNMKKALKLKTTEL 175

RESULT 11
ABB78994
ID ABB78994 standard; Protein; 175 AA.

AC ABB78994;

DT 02-AUG-2002 (first entry)

XX Human XAG-2 protein sequence SEQ ID NO:4473.

XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy.

XX Homo sapiens.

XX W0200229086-A2.

XX 11-APR-2002.

XX 02-OCT-2001; 2001WO-US30732.

XX 02-OCT-2000; 2000US-237271P.

XX (FARB) BAYER CORP.

XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
XX Thilagalingam A, Lewis ME;

XX WPI; 2002-426115/45.

XX N-PSDB; ABQ60777.

XX New isolated nucleic acid that is differentially expressed in cancer
XX tissues useful for determining the presence of colon cancer in a cell
XX or tissue type, and in antisense therapy

XX Claim 5; Fig 3; 796pp; English.

XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
XX expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
XX encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
XX used in antisense therapy. An antibody immunoreactive with a polypeptide
XX encoded by (I) is useful for detecting cancer in a patient sample, and
XX for detecting the presence or absence of a polynucleotide encoded by a
XX nucleic acid which hybridizes to (I) in a cell. A probe/primer derived
XX from (I) can be used for determining the presence of a nucleic acid which

CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists.
XX
XX

SQ Sequence 175 AA;

Query Match 84.6%; Score 899; DB 23; Length 175;
Best Local Similarity 100.0%; Pred. No. 2.2e-90;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 MEKIPVSFAFLLLVALSYTLARDTTVPKAKKDTKDSRPKLPQTLSRGWGDLWTQYEE 91
DB 1 MEKIPVSFAFLLLVALSYTLARDTTVPKAKKDTKDSRPKLPQTLSRGWGDLWTQYEE 60

QY 92 ALYKSTSNKPLMIHHLDECPSHQAALKKVFPAENKEIOKLAEOQVLLNLVYETTDKHLSP 151
DB 61 ALYKSTSNKPLMIHHLDECPSHQAALKKVFPAENKEIOKLAEOQVLLNLVYETTDKHLSP 120

QY 152 DGOYVPRIMFVDPSTLRADITGRYSNRLYAYEPADTALLDDNMKKALKLKTTEL 206
DB 121 DGOYVPRIMFVDPSTLRADITGRYSNRLYAYEPADTALLDDNMKKALKLKTTEL 175

RESULT 12
ABG34043
ID ABG34043 standard; Protein; 175 AA.

XX AC ABG34043;

XX 15-JUL-2002 (first entry)

XX Human Pro peptide #14.

XX Human; PRO: secreted protein; transmembrane protein;
KW genetic disorder; tumour; cancer.

XX Homo sapiens.

XX W0200224888-A2.

XX 28-MAR-2002.

XX 29-AUG-2001; 2001WO-US27099.

XX 01-SEP-2000; 2000US-229896P.

XX 05-SEP-2000; 2000US-230621P.

XX 22-SEP-2000; 2000US-235147P.

XX 10-NOV-2000; 2000WO-US30873.

XX 12-JAN-2001; 2001US-261878P.

XX 16-JAN-2001; 2001US-261910P.

XX 16-JAN-2001; 2001US-261939P.

XX 25-JAN-2001; 2001US-262150P.

XX 25-JAN-2001; 2001US-264395P.

XX 02-FEB-2001; 2001US-266421P.

XX 03-FEB-2001; 2001US-267623P.

XX 28-FEB-2001; 2001WO-US06520.

XX 09-MAR-2001; 2001US-274399P.

XX 03-APR-2001; 2001US-280982P.

XX 04-APR-2001; 2001US-282129P.

XX 09-MAY-2001; 2001US-290589P.

XX 25-MAY-2001; 2001WO-US17092.

XX 01-JUN-2001; 2001WO-US17800.

XX 20-JUN-2001; 2001WO-US19692.

XX 29-JUN-2001; 2001WO-US21066.

XX 09-JUL-2001; 2001WO-US21735.

XX (GETH) GENENTECH INC.

XX PA

XX Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC, Gurney AL, Smith V, Watanabe CK, Wood WI, Zhang Z; Fong S;
XX WPI: 2002-362426/39.
XX N-PSDB; ABK69974.
XX New PRO polypeptides and polynucleotides encoding the polypeptides, useful in gene therapy, chromosome identification, tissue typing, or for genetic analysis of individuals with genetic disorders -
XX Claim 11; Figure 28; 218pp; English.
XX This invention relates to the cDNA and protein sequences of novel secreted and transmembrane polypeptides PRO polypeptides. The invention also comprises a method for producing the proteins of the invention by recombinant means and antibodies specific for the protein of the invention. The antibody may be used for detecting the PRO proteins of the invention and may be used to modify their activity.
XX polynucleotides may be used as hybridisation probes for a cDNA library to isolate the full-length PRO cDNA or to isolate other cDNAs, to construct hybridisation probes for mapping the gene which encodes that PRO and for genetic analysis of individuals with genetic disorders, in assays to identify other proteins or molecules involved in binding reaction, to generate transgenic animals or knock-out animals which turn are useful in the development and screening of therapeutically useful reagents, for chromosome identification, and tissue typing. The PRO polypeptides are useful in gene therapy, and as molecular weight markers for protein electrophoresis purposes. The sequences may also be used to detect overexpression on PRO polypeptides in cancerous tumours and for screening for differentially expressed genes using microarray technology. The present sequence represents a human PRO protein of the invention.

XX Sequence 175 AA;
XX Query Match 84.6%; Score 899; DB 23; Length 175;
XX Best Local Similarity 100.0%; Pred. No. 2.2e-90;
XX Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 32 MEKIPVSAPLLVALSYTLARDTTVPKAKKDTKDSRPKLPQTLSRGWGDLWTQYEE 91
Db 1 MEKIPVSAPLLVALSYTLARDTTVPKAKKDTKDSRPKLPQTLSRGWGDLWTQYEE 60
QY 92 ALYKSTSNKPLMIHLLDECPHSQALKKVFPAENKEIQKLAEQVLLNLVYETTDKHLSP 151
Db 61 ALYKSTSNKPLMIHLLDECPHSQALKKVFPAENKEIQKLAEQVLLNLVYETTDKHLSP 120
QY 152 DGQYVPRIMFVDPSTLVRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 206
Db 121 DGQYVPRIMFVDPSTLVRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 175

RESULT 13
AAU74939
ID AAU74939 standard; Protein; 175 AA.
XX AAU74939;
XX 23-APR-2002 (first entry)
XX Human clone DNA59485-1336 amino acid sequence of PRO1030 protein.
XX Human; clone DNA59482-1336; PRO1030; immune response modulator;
KW PRO polypeptide; immune related disorder; systemic lupus erythematosus;
KW rheumatoid arthritis; osteoarthritis; systemic sclerosis; sarcoidosis;
KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus; anaemia;
KW multiple sclerosis; inflammatory bowel disease; allergic rhinitis.
OS Homo sapiens.
XX Key
XX Location/Qualifiers

FT Peptide 1..20
FT /label= Signal_peptide
FT Protein 21..175
FT /label= Mature_PRO1030_protein
FT Modified-site 30..34
FT /label= Phosphorylation_site
FT /note= "cAMP- and cGMP-dependent protein kinase phosphorylation site"
XX
XX WO200192331-A2.
XX 06-DEC-2001.
XX 25-MAY-2001; 2001WO-US17092.
XX 30-MAY-2000; 2000WO-US14941.
XX 02-JUN-2000; 2000WO-US15264.
XX 01-DEC-2000; 2000WO-US32678.
XX (GETH) GENENTECH INC.
XX Eaton DL, Fong S, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL; Watanabe CK, Wood WI, Zhang Z;
XX WPI: 2002-075461/10.
XX N-PSDB; ABK14011.
XX Isolated PRO221, PRO617, PRO1030, PRO4302 polypeptides, useful for treating immune disorders such as thyroiditis, diabetes mellitus, allergic disease, asthma, allergic rhinitis, atopic dermatitis -
XX Claim 10; Fig 6; 104pp; English.

XX The present invention relates to a new polypeptide having at least 80% amino acid sequence identity to fully defined amino acid sequence of PRO221 (AAU74937), PRO617 (AAU74938), PRO1030 (AAU74939) or PRO4302 (AAU74940) as given in specification. The PRO polypeptides of the invention are useful for treating an immune related disorder such as systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, systemic sclerosis, sarcoidosis, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus, multiple sclerosis, inflammatory bowel disease, allergic rhinitis and anaemia. The methods of the invention are useful for identifying an antisense nucleic acid and for enhancing the infiltration of inflammatory cells into the tissue of a mammal. The anti-PRO antibodies are useful for determining the presence of PRO polypeptides and for diagnosing an immune-related disease in a mammal. The molecules of the invention are useful for identifying a compound that inhibits the activity of PRO polypeptides or expression of a gene encoding the PRO polypeptide. PRO polypeptides are also useful for increasing or decreasing the infiltration of inflammatory cells from the vasculature into a tissue of a mammal, increasing or decreasing activity or proliferation of T-lymphocytes. The present amino acid sequence represents the PRO1030 polypeptide of the invention. The PRO1030 sequence was isolated from the clone DNA59485-1336.

XX Sequence 175 AA;
XX Query Match 84.6%; Score 899; DB 23; Length 175;
XX Best Local Similarity 100.0%; Pred. No. 2.2e-90;
XX Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 32 MEKIPVSAPLLVALSYTLARDTTVPKAKKDTKDSRPKLPQTLSRGWGDLWTQYEE 91
Db 1 MEKIPVSAPLLVALSYTLARDTTVPKAKKDTKDSRPKLPQTLSRGWGDLWTQYEE 60
QY 92 ALYKSTSNKPLMIHLLDECPHSQALKKVFPAENKEIQKLAEQVLLNLVYETTDKHLSP 151
Db 61 ALYKSTSNKPLMIHLLDECPHSQALKKVFPAENKEIQKLAEQVLLNLVYETTDKHLSP 120
QY 152 DGQYVPRIMFVDPSTLVRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 206
Db 121 DGQYVPRIMFVDPSTLVRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 175

CC weight markers and for tissue typing. The PRO nucleic acids have
 CC applications in molecular biology, including use as hybridisation probes,
 CC and in chromosome and gene mapping. AAU83592-AAU83713 represent human PRO
 CC protein sequences of the invention.
 XX
 SQ Sequence 175 AA;
 Query Match 78.1%; Score 830; DB 23; Length 175;
 Best Local Similarity 91.4%; Pred. No. 8.3e-83;
 Matches 160; Conservative 8; Mismatches 7; Indels 0; Gaps 0;
 QY 32 MEKIPVSAPLLVALSYTLARDTVKPGAKKDPKSRPKLPQTLRSRGWQDLIWTQYEE 91
 DB 1 MEKFSVAILLLVAISGTLAKDTTVKSGSKDPKSRPKLPQTLRSRGWQDLIWTQYEE 60
 QY 92 ALYKSKTSNKLPMIITHLDECPHSQALKKVFKAENKEIQKLAEOFVLLNLVYETDKHLSP 151
 DB 61 ALYKSKTSNRPLMVIHLLDECPHSQALKKVFKAENKEIQKLAEOFVLLNLVYETDKHLSP 120
 QY 152 DGQYVPRIMFVDPSTVTRADITGRYSNRLYAYEPADTALLDNMKKALKLLKTEL 206
 DB 121 DGQYVPRIVFVDPSTVTRADITGRYSNRLYAYEPSTALLHDNMKKALKLLKTEL 175
 RESULT 15
 AAG00109
 ID AAG00109 standard; Protein; 132 AA.
 XX
 AC AAG00109;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 4190.
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX Homo sapiens.
 OS EP1033401-A2.
 PN 06-SEP-2000.
 PD 21-FEB-2000; 2000EP-0200610.
 PF 26-FEB-1999; 99US-0122487.
 PR (GEST) GENSET.
 PA Dumas Milne Edwards J, Duclert A, Giordano J;
 PI WPI: 2000-500381/45.
 DR N-PSDB; AAC00115.
 XX
 XX The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dr primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.

CC weight markers and for tissue typing. The PRO nucleic acids have
 CC applications in molecular biology, including use as hybridisation probes,
 CC and in chromosome and gene mapping. AAU83592-AAU83713 represent human PRO
 CC protein sequences of the invention.
 XX
 SQ Sequence 175 AA;
 Query Match 78.1%; Score 830; DB 23; Length 175;
 Best Local Similarity 91.4%; Pred. No. 8.3e-83;
 Matches 160; Conservative 8; Mismatches 7; Indels 0; Gaps 0;
 QY 32 MEKIPVSAPLLVALSYTLARDTVKPGAKKDPKSRPKLPQTLRSRGWQDLIWTQYEE 91
 DB 1 MEKFSVAILLLVAISGTLAKDTTVKSGSKDPKSRPKLPQTLRSRGWQDLIWTQYEE 60
 QY 92 ALYKSKTSNKLPMIITHLDECPHSQALKKVFKAENKEIQKLAEOFVLLNLVYETDKHLSP 151
 DB 61 ALYKSKTSNRPLMVIHLLDECPHSQALKKVFKAENKEIQKLAEOFVLLNLVYETDKHLSP 120
 QY 152 DGQYVPRIMFVDPSTVTRADITGRYSNRLYAYEPADTALLDNMKKALKLLKTEL 206
 DB 121 DGQYVPRIVFVDPSTVTRADITGRYSNRLYAYEPSTALLHDNMKKALKLLKTEL 175
 RESULT 14
 AAU83628
 ID AAU83628 standard; Protein; 175 AA.
 XX
 AC AAU83628;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE Human PRO protein, Seq ID No 74.
 XX
 KW Human; secreted protein; PRO; tumour; lung cancer; colon cancer;
 KW breast cancer; prostate tumour; rectal tumour; liver tumour;
 KW pericyte cell proliferation; chondrocyte cell proliferation;
 KW tumour necrosis factor-alpha.
 XX Homo sapiens.
 OS WO200208288-A2.
 PN 31-JAN-2002.
 PD 29-JUN-2001; 2001WO-US21066.
 PF 20-JUL-2000; 2000US-219556P.
 PR 25-JUL-2000; 2000US-220585P.
 PR 25-JUL-2000; 2000US-220605P.
 PR 25-JUL-2000; 2000US-220607P.
 PR 25-JUL-2000; 2000US-220624P.
 PR 25-JUL-2000; 2000US-220638P.
 PR 25-JUL-2000; 2000US-220664P.
 PR 25-JUL-2000; 2000US-220666P.
 PR 26-JUL-2000; 2000US-220893P.
 PR 28-JUL-2000; 2000WO-US20710.
 PR 23-AUG-2000; 2000WO-US23522.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 15-SEP-2000; 2000US-000000P.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 28-NOV-2000; 2000US-253646P.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 20-DEC-2000; 2000US-0747259.
 PR 20-DEC-2000; 2000WO-US34956.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 10-MAY-2001; 2001US-0854280.
 PR 25-MAY-2001; 2001WO-US17092.
 XX (GETH) GENENTECH INC.
 PA Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
 PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
 XX WPI: 2002-172001/22.
 DR N-PSDB; ABK33572.
 XX
 XX One hundred and twenty two nucleic acids encoding PRO polypeptides,
 CC useful for treating a PRO related disorder and for diagnosing tumours
 CC such as lung cancer, colon cancer, breast tumour, prostate tumour, rectal
 CC tumour or liver tumour -
 XX
 XX Claim 11; Figure 74; 359pp; English.
 CC The invention relates to one hundred and twenty two nucleic acids
 CC encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides
 CC encode human secreted proteins. The PRO nucleic acids, polypeptides,
 CC agonists and antagonists are useful for treating a PRO related disorder.
 CC The PRO polypeptides are useful for diagnosing tumours, especially lung
 CC cancer, colon cancer, breast tumour, prostate tumour, rectal tumour or
 CC liver tumour. The PRO polypeptides are useful for stimulating the
 CC proliferation of, or gene expression, in pericyte cells, for stimulating
 CC the proliferation or differentiation of chondrocyte cells, for
 CC stimulating the release of tumour necrosis factor-alpha from human blood,
 CC for stimulating or inhibiting the proliferation of normal human dermal
 CC fibroblast cells. The PRO polypeptide may also be used as molecular

```

xx
SQ Sequence 132 AA;
Query Match 61.7%; Score 656; DB 21; Length 132;
Best Local Similarity 100.0%; Pred. No. 7.2e-64;
Matches 126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 32 MEKIPVSFLLVLSYTLARDTTVKPGAKKDTKDSRPKLPQTLSRGWGDLIWTQTYEE 91
Db 1 MEKIPVSFLLVLSYTLARDTTVKPGAKKDTKDSRPKLPQTLSRGWGDLIWTQTYEE 60
Qy 92 ALYKSKTSNKPLMLIIHHLDECPHSQALKKVFPAENKEIQKLAEOFLNLNLYETTDKHLSP 151
Db 61 ALYKSKTSNKPLMLIIHHLDECPHSQALKKVFPAENKEIQKLAEOFLNLNLYETTDKHLSP 120
Qy 152 DGQYVP 157
Db 121 DGQYVP 126

```

Search completed: January 17, 2003, 12:55:56
Job time : 37 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 12:55:16 ; Search time 20 Seconds
(without alignments)
990.185 Million cell updates/sec

Title: US-09-646-778A-288
Perfect score: 1063
Sequence: 1 RLSCAGTSLSGSPHPSRLT.....DTALLDNMKALKLLKTEL 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 73:*

1: Piri:*

2: Piri:*

3: Piri:*

4: Piri:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	899	84.6	175	2 JE0350	Anterior gradient-
2	199	18.7	186	2 T31643	hypothetical prote
3	119	11.2	150	2 D69100	probable protein d
4	87.5	8.2	458	2 S50919	hypothetical prote
5	87	8.2	348	2 G72016	probable disulfide
6	87	8.2	348	2 C86607	disulfide bond iso
7	85.5	8.0	1830	1 S31988	myosin-v - chicken
8	84.5	7.9	1151	2 A61717	probable peptidogl
9	84	7.9	771	2 S51421	hypothetical prote
10	83.5	7.9	217	2 T37859	probable transcrip
11	83.5	7.9	918	2 C96829	unknown protein F1
12	83	7.8	479	2 T40683	cell cycle protein
13	83	7.8	507	2 S56143	cell cycle protein
14	82.5	7.8	788	2 C70441	lipocate-protein li
15	81.5	7.7	893	2 G59431	phosphatidylinosit
16	81.5	7.7	970	2 S29069	inositol polyphosp
17	81.5	7.7	1158	2 T50454	probable rho1 GDP-
18	81.5	7.7	1397	2 T10466	DNA topoisomerase
19	81	7.6	610	2 T22909	hypothetical prote
20	81	7.6	635	2 T23874	hypothetical prote
21	80	7.5	154	2 F82633	probable protein d
22	80	7.5	259	2 A96740	F14023.14 (importe
23	80	7.5	843	2 S53947	hypothetical prote
24	79.5	7.5	277	2 T37629	hypothetical prote
25	79.5	7.5	301	2 G71929	GTP-binding protei
26	79.5	7.5	642	2 F83718	ABC transporter (A
27	79.5	7.5	921	2 F84593	hypothetical prote
28	79	7.4	477	2 T37791	probable rna polym
29	79	7.4	510	1 S43516	carboxypeptidase C

30	78.5	7.4	280	2 A69590	shikimate 5-dehydr
31	78.5	7.4	917	2 T05430	hypothetical prote
32	78.5	7.4	1033	2 A96714	hypothetical prote
33	78.5	7.4	1262	2 I48855	tripeptidyl-peptid
34	78.5	7.4	1570	2 AC2012	hypothetical prote
35	78	7.3	337	2 S06956	segmentation prote
36	78	7.3	1828	2 B59254	myosin heavy chain
37	78	7.3	1853	1 A46761	myosin heavy chain
38	78	7.3	1855	2 A59254	formate dehydrogen
39	76.5	7.2	985	2 E69850	RNA-directed DNA p
40	76.5	7.2	1025	2 T10259	thioredoxin disulf
41	76	7.1	166	2 D72018	thioredoxin disulf
42	76	7.1	166	2 D86606	probable uracil ki
43	76	7.1	367	2 T39574	oligopeptide-bindi
44	76	7.1	543	1 F64871	probable peptide A
45	76	7.1	583	2 E75529	

ALIGNMENTS

RESULT 1

JE0350

Anterior gradient-2 - human

C:Species: Homo sapiens (man)

C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 21-Jul-2000

C:Accession: JE0350

R:Thompson, D.A.; Weigel, R.J. 111-116, 1998

Biochem. Biophys. Res. Commun. 251, 111-116, 1998
A:Title: hAG-2, the human homologue of the Xenopus laevis cement gland gene XAG-2, is

A:Reference number: JE0350; MUID:99009231; PMID:9790916

A:Accession: JE0350

A:Molecule type: mRNA

A:Residues: 1-175 <THO>

A:CROSS-references: GB:AF007791; NID:g3779196; PIDN:AAC77358.1; PID:g3779197
C:Comment: This protein is coexpressed with estrogen receptor (ER) in breast cancer c

C:Genetics:

A:Gene: hAG-2

C:Superfamily: human anterior gradient-2

Query Match 84.6%; Score 899; DB 2; Length 175;

Best Local Similarity 100.0%; Pred. No. 3.1e-74;

Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 MEKIPVSAPLLVALSYTLARDTVKPGAKKDKSRPKLPQTLSPGWDQLWTQYEE 91

Db 1 MEKIPVSAPLLVALSYTLARDTVKPGAKKDKSRPKLPQTLSPGWDQLWTQYEE 60

Qy 92 ALYKSKTSNKPLMIHHLDCEPHSQALKKVFENKEIOKLAEQFVLLMLVYETTDKHLSP 151

Db 61 ALYKSKTSNKPLMIHHLDCEPHSQALKKVFENKEIOKLAEQFVLLMLVYETTDKHLSP 120

Qy 152 DQYVPRIMFVDPSTVRADITGRYSNRLYAYEPADTALLDNMKALKLLKTEL 206

Db 121 DQYVPRIMFVDPSTVRADITGRYSNRLYAYEPADTALLDNMKALKLLKTEL 175

RESULT 2

T31643

hypothetical protein Y57A10A.u - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000

C:Accession: T31643

R:Smyle, R. submitted to the EMBL Data Library, September 1999

A:Reference number: Z21048

A:Accession: T31643

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-186 <WIL>

A:CROSS-references: EMBL:AL117195; PIDN:CAB55026.1; CESP:Y57A10A.u

A:Experimental source: clone Y57A10A

C:Genetics:

A:Gene: CESP:Y57A10A.u

A:Introns: 70/1

C:Superfamily: human anterior gradient-2

Query Match 18.7%; Score 199; DB 2; Length 186;
Best Local Similarity 30.0%; Pred. No. 1.1e-10;
Matches 51; Conservative 38; Mismatches 57; Indels 24; Gaps 6;

QY 41 LLLVALSYTLARDTTVPKAKKDTKDSRPKLPQTLSRGWGDOLITQTYEEALYKSKTSN 100

Db 4 LLLVAL-----VSASAYASFDKIQDSIQNPLARGFGDDIAWK-WEDIAETALDFTD 53

QY 101 KPLMIHHLDPCPSQALKKFAEN---KEIQKLAQFVLLNLV--YETDKHLSPDQGY 155

Db 54 KPFIILLHKSCHACKALKKTFQSQNAKKAFFKLSERFVWNTEDDDPFEERYPDGKY 113

QY 156 VPRIMFVDPSTLVRADITGRYSNRL-----YAYEPADTALLLDNKKALK 200

Db 114 IPRULFDKN-----GDLQEPKNAKAEYKKNYAYYSPPADILNSMKDVLK 159

RESULT 3

D69100

probable protein disulfide-isomerase (EC 5.3.4.1) MTH1745 [similarity] - Methanobacterium

C:Species: Methanobacterium thermoautotrophicum

C>Date: 03-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Sep-2000

C:Accession: D69100

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicalre, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N. K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct

A:Reference number: A69000; MUID:98037514; PMID:9371463

A:Accession: D69100

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-150 <MTH>

A:Cross-references: GB:AE000930; GB:AE000666; NID:g2622872; PIDN:AAB86215.1; PID:g262287

A:Experimental source: strain Delta H

C:Genetics:

C:Keywords: intramolecular oxidoreductase; isomerase

Query Match 11.2%; Score 119; DB 2; Length 150;
Best Local Similarity 27.1%; Pred. No. 0.0015;
Matches 35; Conservative 28; Mismatches 46; Indels 20; Gaps 4;

QY 39 AFLLVALSYTLARDTTVPKAKKDTKDSRPKLPQTLSRGWGDOLITQTYEEALYKSKT 98

Db 19 AGLSLVLIITV-----QPRVPSLTDEKO-LKWYTEHDEAIKEASR 60

QY 99 SNKPLMIHHLDPCPSQALKKFAENKEIQ-KLAQFVLLNLVYETDKHLSPDQGY-V 156

Db 61 TGNKPMFVSASWCPACOKLESETLQNTQVORLAEFDIAVKIDVTSPPALSSRYIYGV 120

QY 157 PRIMFVDPSS 165

Db 121 PTVIIDPS 129

RESULT 4

S50919

hypothetical protein YMR191w - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YMR191w

C:Species: Saccharomyces cerevisiae

C>Date: 10-Feb-1995 #sequence_revision 12-May-1995 #text_change 19-Apr-2002

C:Accession: S50919; S55721

R:Pearson, D.; Bowman, S.

submitted to the EMBL Data Library, January 1995

A:Reference number: S50919

A:Accession: S50919

A:Molecule type: DNA

A:Residues: 1-358 <PEA>

A:Cross-references: EMBL:Z47815; NID:g642280; PIDN:CAA87812.1; PID:g642283; MIPS:YMR1
R:Gangloff, S.; McDonald, J.P.; Bendixen, C.; Arthur, L.; Rothstein, R.
Mol. Cell. Biol. 14, 8391-8398, 1994

A:Title: The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA helicase homo

A:Reference number: A56359; MUID:95059068; PMID:7969174

A:Accession: S55721

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-127, Y, 129-131, LALG, 136-249 <GNA>

A:Cross-references: EMBL:U22341; NID:g726277; PIDN:AAB60288.1; PID:g726278

A:Note: The nucleotide sequence was submitted to the EMBL Data Library, March 1995

C:Genetics:

A:Cross-references: SGD:S0004803

A:Map position: 13R

C:Superfamily: Saccharomyces cerevisiae hypothetical protein YMR191w

Query Match 8.2%; Score 87.5; DB 2; Length 458;

Best Local Similarity 20.1%; Pred. No. 4.5;

Matches 55; Conservative 35; Mismatches 74; Indels 109; Gaps 13;

QY 15 PSRLTQ-----RWVRSRVAMEKIPVSAPLLIVALS----- 47

Db 82 PTRNAVGVGNWGMWLMRMSVHLRQITKSLDRLISLHSGNFHSQYRNRIEVTWMSLFE 141

QY 48 -----YTLARDTTVPKAKKDTK-----DSRPKLPOTLSRG-----GD 81

Db 142 ASTAFRRASGLTVSLTRGIAEDHFRPVPVSKFASFPKAPGRLFTNNMTTSK 201

QY 82 QLWITQTYEALYK-----SKTSNKLMT-----IHHLECPHSQ----- 116

Db 202 RLIGQRAYSTSIKFTQEAANNNTISLRCFFNSLGLNOCSSCKAYQNASVTSKQD 261

QY 117 -----ALKV-----FAENKEIQKLAQFVLLNLVYETDKHLSPDQGY----- 156

Db 262 HVQPVALKKLSQKIDNIRNLELFKMK--TQNEVDETSAIYMEKPGSIEFTISEFNV 319

QY 157 -----PRIMFVDPSTLVRADITGRYSNRLYAYE 184

Db 320 NGTFSAPIFLDPSLL--ADLDEMIRN--YKYE 348

RESULT 5

G72016

probable disulfide bond isomerase - Chlamydia pneumoniae (strains CWL029 and AR39

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000

C:Accession: G72016; B81523

R:Kelman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,

Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; MUID:99206606; PMID:10192388

A:Accession: G72016

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-348 <ARN>

A:Cross-references: GB:AE001673; GB:AE001363; NID:g437752; PIDN:AAD19071.1; PID:g437

A:Experimental source: strain CWL029

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke

, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39

A:Reference number: A81500; MUID:20150255; PMID:10684935

A:Accession: B81523

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-348 <REA>

A:Cross-references: GB:AE002251; GB:AE002161; NID:g7189833; PIDN:AAF38711.1; PID:g718

A:Experimental source: strain AR39, HL cells

C:Genetics:

A:Gene: Cpn0933; Cpn0928

Query Match 8.2%; Score 87; DB 2; Length 348;

Best Local Similarity 25.0%; Pred. No. 3.5;

[illegible]

C:Genetics: SPDB:SPBC776.12c
A:Gene: SPDB:SPBC776.12c
A:Map position: 2
A:Introns: 77/1; 92/2; 331/1
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

Query Match 7.8%; Score 83; DB 2; Length 479;
Best Local Similarity 22.4%; Pred. No. 12;
Matches 28; Conservative 24; Mismatches 55; Indels 18; Gaps 5;

QY 83 LIWQTYEEALYKSTNKP-----LMIHHLDECPHSQALKKVFPAENKEIOKLAEOFVLL 138
DB 352 ILWASCSGASIIYKEKLRHKPSQERLCDFLEKCLDCKNKRISAE-----EALDHDFLYL 407
QY 139 -NLVYETTDKHLSPDQGVYPRIMFVDPSTLVTRADITGRYSNRLYAY---EPADTALLDN 194
DB 408 DNLAYEKDDDDTAFDNS-----FGETSFEKDEDLTAKHLSHILDFKEQETDEPTSLSK 461
QY 195 MKKAL 199
DB 462 RKRSI 466

RESULT 13
S56143
C:Species: Schizosaccharomyces pombe
C:Date: 28-Oct-1995 #sequence_revision 10-Nov-1995 #text_change 10-Dec-1999
C:Accession: S56143; T40682
R:Masai, H.; Miyake, T.; Arai, K.
EMBO J. 14, 3094-3104, 1995
A:Title: hsk1(+), a Schizosaccharomyces pombe gene related to Saccharomyces cerevisiae C
A:Reference number: S56143; MUID:95347336; PMID:7621824
A:Accession: S56143
A:Molecule type: DNA
A:Residues: 1-507 <MAS>
A:Cross-references: GB:D50493; NID:g807845; PIDN:BAA09087.1; PID:d1009722; PID:g807846
A:Experimental source: strain jy2
R:Lyne, M.; Rajandream, M.A.; Barrall, B.G.; Wedler, H.; Kutzner, M.; Wambutt, R.
submitted to the EMBL Data Library, January 1999
A:Reference number: 221944
A:Accession: T40682
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-507 <LYN>
A:Cross-references: EMBL:AL035263; PIDN:CAA22885.1; GSPDB:GN00067; SPDB:SPBC776.12c
A:Experimental source: strain 972h-; cosmid c776
C:Genetics:
A:Gene: hsk1; SPBC776.12c
A:Map position: 2
A:Introns: 77/1; 359/1
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: phosphotransferase; protein kinase
F:66-404/Domain: protein kinase homology <KIN>

Query Match 7.8%; Score 83; DB 2; Length 507;
Best Local Similarity 22.4%; Pred. No. 13;
Matches 28; Conservative 24; Mismatches 55; Indels 18; Gaps 5;

QY 83 LIWQTYEEALYKSTNKP-----LMIHHLDECPHSQALKKVFPAENKEIOKLAEOFVLL 138
DB 380 ILWASCSGASIIYKEKLRHKPSQERLCDFLEKCLDCKNKRISAE-----EALDHDFLYL 435
QY 139 -NLVYETTDKHLSPDQGVYPRIMFVDPSTLVTRADITGRYSNRLYAY---EPADTALLDN 194
DB 436 DNLAYEKDDDDTAFDNS-----FGETSFEKDEDLTAKHLSHILDFKEQETDEPTSLSK 489
QY 195 MKKAL 199
DB 490 RKRSI 494

RESULT 14

C70441

lipoate-protein ligase A - Aquifex aeolicus.
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: C70441
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: C70441
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-788 <AQF>
A:Cross-references: GB:AE000748; NID:g2983960; PIDN:AAC07512.1; PID:g2983965; GB:AE00
A:Experimental source: strain Vfs
C:Genetics:
A:Gene: lplA

Query Match 7.8%; Score 82.5; DB 2; Length 788;
Best Local Similarity 23.0%; Pred. No. 25;
Matches 49; Conservative 34; Mismatches 73; Indels 57; Gaps 10;

QY 25 VRKSRVAMEKIPVSFAFLLLVALSYTLIARDTTVKPGAKDKTSR-PKLPTLSR-GW--G 80
DB 585 VQRSAQKQPVENENLKEFIPQYKVPK---VAPGQEDM--SRLPIVRETSTEIGFING 639
QY 81 DQLIWQTYEEALYKSTNKPIMIIHHLDECPHSQALKKVFPAENKEIOKLAEOFVLLNL 140
DB 640 QKVSEALQVAVELLSKAKRPTLVGPL-----VLWANEAKKKEGIIRKLKE 689
QY 141 VYETTDKHLSPDQGVYPRIMFVD-----PSLTV-----RADITGR 175
DB 690 KYPNLNIHLPLD--YKPKNKKVDFSEVDPNPHLSVLGYEHDITLMIGVHCYRTDFTIR 747
QY 176 YSNR-----LYAYEPADTALLDNMKK 197
DB 748 MLKKYTDTKIVTLCTLYGHPDADVSLAGVNSKK 780

RESULT 15
G59431
phosphatidylinositol polyphosphate 5-phosphatase, isoform b [imported] - huamn
C:Species: Homo sapiens (man)
C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 19-Jul-2002
C:Accession: G59431; H59431; A59443; B59432; C59432; D59432; E59432; F59432
R:Duden, R.; Griffiths, G.; Frank, R.; Argos, P.; Kreis, T.E.
Cell 64, 649-665, 1991
A:Title: Beta-COP, a 110 kd protein associated with non-clathrin-coated vesicles and
A:Reference number: G59431; MUID:91121521; PMID:1840503
A:Accession: G59431
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-893 <DUD>
A:Cross-references: GB:NP_001578; PID:g13325070; PIDN:NP_001578.2
R:Mueller, O.T.; Hartsfield, J.K. Jr.; Gallardo, L.A.; Essig, Y.P.; Miller, K.L.; Pap
Am. J. Hum. Genet. 49, 804-810, 1991
A:Title: Lowe oculocerebrorenal syndrome in a female with a balanced X;20 translocati
A:Reference number: H59431
A:Accession: H59431
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-893 <MUE>
A:Cross-references: GB:NP_001578; PID:g13325070; PIDN:NP_001578.2
R:Attree, O.; Olivos, I.M.; Okabe, I.; Bailey, L.C.; Nelson, D.L.; Lewis, R.A.; McInn
Nature 358, 239-242, 1992
A:Title: The Lowe's oculocerebrorenal syndrome gene encodes a protein highly homologo
A:Reference number: A59443; MUID:92334430; PMID:1321346
A:Accession: A59443
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-893 <ATT>
A:Cross-references: GB:NP_001578; PID:g13325070; PIDN:NP_001578.2

Search completed: January 17, 2003, 12:57:23
Job time : 22 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 12:55:16 ; Search time 12 Seconds
(without alignments)
712.010 Million cell updates/sec

Title: US-09-646-778A-288
Perfect score: 1063
Sequence: 1 RLSCAGTILSGSPHPSRLT.....DTALLLDNMKALKLLKTEL 206

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	448	42.1	183	1 XAG_XENLA	P55868 xenopus lae
2	437.5	41.2	185	1 NP77_XENLA	P55869 xenopus lae
3	87.5	8.2	458	1 YMS1_YEAST	P42933 saccharomyc
4	85.5	8.0	1829	1 MYSA_CHICK	Q02440 gallus gall
5	84	7.9	811	1 RFXL_YEAST	P48743 saccharomyc
6	83.5	7.9	217	1 YD67_SCHPO	Q10319 schizosacch
7	83	7.8	507	1 HSK1_SCHPO	P50582 schizosacch
8	81.5	7.7	1398	1 OCRL_HUMAN	Q01968 homo sapien
9	81.5	7.7	1398	1 TOP2_PLAFA	P41001 plasmodium
10	80	7.5	339	1 YTH6_RHOER	P43484 rhodococcus
11	80	7.5	843	1 YMS1_YEAST	Q05050 saccharomyc
12	79.5	7.5	301	1 ERA_HELPJ	Q921w0 helicobacte
13	79	7.4	510	1 CBPI_ORISA	P37890 oryza sativ
14	78.5	7.4	280	1 AROE_BACSU	P54374 bacillus su
15	78.5	7.4	1262	1 TPP2_MOUSE	Q64514 mus musculu
16	78.5	7.4	1828	1 MYSA_RAT	Q9qvf3 rattus norv
17	78	7.3	337	1 HAIR_DROME	P14003 drosophila
18	78	7.3	887	1 ITH3_RAT	Q63416 rattus norv
19	78	7.3	1853	1 MYSA_MOUSE	Q99104 mus musculu
20	78	7.3	1855	1 MYSA_HUMAN	Q9y411 homo sapien
21	77	7.2	729	1 YP65_CAEEL	Q09214 caenorhabdi
22	76.5	7.2	1025	1 PO21_NASVI	Q03278 nasonia vit
23	76	7.1	543	1 OPPI_ECOLI	P23843 escherichia
24	76	7.1	2145	1 U520_CAEEL	Q9u2d0 caenorhabdi
25	75.5	7.1	580	1 TREZ_MYCTO	Q10769 mycobacteri
26	75.5	7.1	2318	1 NTC3_MOUSE	Q61982 mus musculu
27	75	7.1	1694	1 SN_MOUSE	Q62230 mus musculu
28	74.5	7.0	381	1 PRLP_BOVIN	Q9qkn8 bos taurus
29	74.5	7.0	905	1 ZO3_MOUSE	Q9qxl1 mus musculu
30	74.5	7.0	2319	1 NTC3_RAT	Q9rl72 rattus norv
31	74.5	7.0	2321	1 NTC3_HUMAN	Q9um47 homo sapien
32	74	7.0	260	1 TATD_ECOLI	P27859 escherichia
33	74	7.0	986	1 GM13_RAT	Q62839 rattus norv

34	74	7.0	1034	1 POL_HV2CA	P24107 human immun
35	74	7.0	1651	1 VIT6_CAEEL	P18948 caenorhabdi
36	73.5	6.9	233	1 YEG7_YEAST	P39983 saccharomyc
37	73.5	6.9	695	1 TRT_PICST	P34736 pichia stip
38	73.5	6.9	971	1 CSEL_PAGMA	Q9ptu3 pagrus majo
39	73	6.9	316	1 COAA_YERPE	Q8zan6 yersinia pe
40	73	6.9	448	1 GUN_CLOSA	P15704 clostridium
41	73	6.9	522	1 UNC7_CAEEL	Q03412 caenorhabdi
42	72.5	6.8	177	1 YP98_CAEEL	Q09246 caenorhabdi
43	72.5	6.8	188	1 KTHY_METJA	Q57741 methanococ
44	72.5	6.8	689	1 YVAL_BACSU	P37512 bacillus su
45	72.5	6.8	823	1 DMSA_RHOCA	Q52675 rhodobacter

ALIGNMENTS

RESULT 1					
XAG_XENLA					
ID	XAG_XENLA	STANDARD;	PRT;	183 AA.	
AC	P55868;				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DE	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	Putative secreted protein XAG precursor.				
GN	XAG.				
OS	Xenopus laevis (African clawed frog).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;				
OC	Xenopodinae; Xenopus.				
NCBI_TaxID=8355;					
OR	NCBI				
OR	SEQUENCE FROM N.A.				
RP	TISSUE=Cement gland.				
RC	MEDLINE=97003224; PubMed=88050563;				
RA	Sive H., Bradley L.;				
RT	"A sticky problem: the Xenopus cement gland as a paradigm for				
RT	anteposterior patterning."				
RL	Dev. Dyn. 205:265-280 (1996).				
CC	- SUBCELLULAR LOCATION: Secreted (Probable).				
CC	- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVEL IN CEMENT GLAND.				
CC	- SIMILARITY: HIGH; TO XENOPUS NP77.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; U76752; AAB18819.1; -				
KW	Signal.				
FT	SIGNAL. 1 18 POTENTIAL.				
FT	CHAIN 19 183 PUTATIVE SECRETED PROTEIN XAG.				
SQ	SEQUENCE 183 AA; 20475 MW; 95E12B06FCB8DCE8 CRC64;				
Query Match	42.1%; Score 448; DB 1; Length 183;				
Best Local Similarity	49.4%; Pred. No. 8e-35;				
Matches	88; Conservative 29; Mismatches 51; Indels 10; Gaps 3;				
QY	39 AFLLLVALSYTLARDITVKP---GAKKDTKDSRPKLP-----QTLGRGWGDLIWTQTY 89				
DB	6 SLVCLVLLCSALGEAVLKKPKKQAGTDTKTQDEPAPIKTKGLTKLDRNGESIEWQTY 65				
QY	90 EEAFLYKSTNSPLMIHHLDCEPHSQALKKVFPAENKEIOKLA-EQVVLNLNYETTDKH 148				
DB	66 EEGLAKARENKPLMVIHHLEDCPYSTALKAFVPAENMAQKLAQEDFIMLNHPVADEN 125				
QY	149 LSPDGQYVPRIMEVDPSTLVTRADITGRYSNRLYAYEPADTALLLDNMKALKLLKTEL 206				
DB	126 QSPDGHVPRVIFDPSLTVRSOLKGRYGNKMYAYDADDIPELITNMKAKSFLKTEL 193				

```

RESULT 2
NP77_XENLA          STANDARD;          PRT;    185 AA.
ID  NP77_XENLA
AC  P55869;
DT  01-NOV-1997 (Rel. 35, Created)
DT  15-JUL-1998 (Rel. 36, Last sequence update)
DT  15-JUL-1998 (Rel. 36, Last annotation update)
DE  Putative secreted protein NP77 precursor.
GN  NP77.
OS  Xenopus laevis (African clawed frog).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC  Xenopodinae; Xenopus.
OX  NCBI_TaxID=8355;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Cement gland;
RA  Aberger F., Schuren C., Lepperdinger G., Richter K., Grunz H.;
RL  Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
CC  -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC  -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVEL IN CEMENT GLAND
CC  (POTENTIAL).
CC  -!- SIMILARITY: HIGH, TO XENOPUS XAG.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; J082110; AAB49974.1; -
KW  Signal.
FT  SIGNAL.
FT  CHAIN.
SQ  SEQUENCE 185 AA; 20442 MW; AE3807C926044509 CRC64;

Query Match      41.2%; Score 437.5; DB 1; Length 185;
Best Local Similarity 53.5%; Pred. No. 7.7e-34;
Matches 84; Conservative 26; Mismatches 46; Indels 1; Gaps 1;

QY  51  ARDTTVKPKAKKTDKSRPKPLQTLRGWDQLIWTQTYEEALVKSNTKPLMIHHL 110
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  29  AARTDTNGAKSEPAVPKTKGLTLDRGWGEDIWTAQTYEEGLAKARENKPLMVIH 88
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  111 ECPHSQLKKVFAENKEIQKLA-EQFVLLNLVYETTDKHLSPDQGVYPRIMFVDP 169
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  89  DCPYSIALKKAFVADKMAQKLAQEDFIMLNLVHPVADENQSPDGHYVPKGIFID 148
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  170 ADITGRYSNRLVAYEPADTALLDNNKKALKLLKTEL 206
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  149  SLDKGRYGNKLYAYDADDIPELITNNMKKAKSFLKTEL 185
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 3
YM51_YEAST
ID  YM51_YEAST
AC  P42933;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  01-NOV-1997 (Rel. 35, Last annotation update)
DE  Hypothetical 52.2 kDa protein in SGI1-MRPL24 intergenic region.
GN  YMR191W OR YMR646.03.
OS  Saccharomyces cerevisiae (Baker's yeast).
OC  Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC  Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX  NCBI_TaxID=4932;
RN  [14]
RP  SEQUENCE FROM N.A.
RC  STRAIN=S288c / AB972;
RA  Pearson P., Bowman S., Barrell B.G., Rajandream M.A.;

```

```

RL  Submitted (JAN-1995) to the EMBL/GenBank/DDBJ databases.
RN  [2]
RP  SEQUENCE OF 1-249 FROM N.A.
RC  STRAIN=W303;
RX  MEDLINE=95059068; PubMed=7969174;
RA  Gangloff S., McDonald J.P., Bendixen C., Arthur L., Rothstein R.;
RT  "The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA
RT  helicase homolog: a potential eukaryotic reverse gyrase.";
RL  Mol. Cell. Biol. 14:8391-8398(1994).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; Z47815; CAA87812.1; -
DR  EMBL; U22341; AAB60288.1; -
DR  SGD; S0004803; YMR191W.
KW  Hypothetical protein.
FT  CONFLICT 128 128      N -> Y (IN REF. 2).
FT  CONFLICT 132 135      FVTW -> LALG (IN REF. 2).
SQ  SEQUENCE 458 AA; 52183 MW; C31107D7B28DAE33 CRC64;

Query Match      8.2%; Score 87.5; DB 1; Length 458;
Best Local Similarity 20.1%; Pred. No. 1.3;
Matches 55; Conservative 35; Mismatches 74; Indels 109; Gaps 13;

QY  15  PSRLRTQG-----RMVRSRVAMKIPYSAFLLLVALS----- 47
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  82  PTRNNAVGNNMWSMRLSRVHLRQITKSLDRLISLHSHGNSHQYNRNFVTTWKS 141
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  48  ----YTLARDTTVPKGAKKDK-----DSRPKLPQTLRG-----W----GD 81
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  142  ASTAFRRASGLTVSLTRGRIAREDFHFRVPVNSKFAFPVPKGAAPRGLTNNW 201
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  82  QLIWTQTYEEALYK---SKTSNKNPLMI-----IHLDCEPHSQ----- 116
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  202  RLLGORAYTSIKFTQPAVNMTISLRCFFNSLGLNQCSHNSCKAYONASNVTS 261
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  117  ----ALKKV-----FAENKEIQKLAEQFVLLNLVYETTDKHLSPDQGVY----- 156
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  262  HVQPVALKKLSQKDFINFLNLELFKIMK--TQNEVVDTSAYYMEKPGSVIEFT 319
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  157  ----PRMFVDPSTVTRADITGRYSNRLIAYE 184
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  320  NGTESAPLSFLDPSLL--ADLDEMIRN--YKYE 348
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
MY5A_CHICK
ID  MY5A_CHICK
AC  Q02440;
DT  01-JUN-1994 (Rel. 29, Created)
DT  01-JUN-1994 (Rel. 29, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Myosin Va (Myosin 5A) (Dilute myosin heavy chain, non-muscle) (Myosin
DE  heavy chain p190) (Myosin-V).
GN  MYO5A.
OS  Gallus gallus (Chicken).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC  Gallus.
OX  NCBI_TaxID=9031;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Brain;
RX  MEDLINE=93012002; PubMed=1383040;
RA  Sanders G., Lichte B., Meyer H.E., Kilmann M.W.;
RT  "cDNA encoding the chicken ortholog of the mouse dilute gene product.
RT  Sequence comparison reveals a myosin I subfamily with conserved C-

```


RESULT 6
 DB 77 SPDP -EVETGCGEDIMVRIF-----FAPEAHEKALTGWHHLKHPYGRMEEMKASG 130
 QY 127 EIQLAEQVLLNLYETTDKHLSP-----DGQYVPRIMFVDPSTVRADITGRYSNRLYA 182
 DB 131 LVESVQVEEIVNEPEYTYKLLSQNPIDGHH-----GLAVESEDPHFPSQOLEQ 180
 QY 183 YEPADTALLDNKMKALKLKTTEL 206
 DB 181 DEADKLDFAIQEVKVTIEMKQOV 204
 RESULT 7
 HSK1_SCHPO
 ID HSK1_SCHPO STANDARD: PRT: 507 AA.
 AC P50582; 094678;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cell cycle protein kinase hsk1 (EC 2.7.1.-)
 GN HSK1 OR SPBC776.12C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 CX NCBI_TaxID=4896;
 [1]
 RN SEQUENCE FROM N.A. (LONG ISOFORM).
 RP STRAIN=JY2;
 RC MEDLINE=95347336; PubMed=7621824;
 RA Masai H., Miyake T., Arai K.-I.;
 RT "hsk1+, a Schizosaccharomycetes pombe gene related to Saccharomyces
 RL cerevisiae CDC7, is required for chromosomal replication.";
 RN EMBO J. 14:3094-3104(1995).
 [2]
 RN SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RP STRAIN=972;
 RC MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jageis K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford J., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Benito J.,
 RA Lucas M., Rochet M., Gallard C., Talleda V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RL "The genome sequence of Schizosaccharomycetes pombe.";
 CC Nature 415:871-880(2002).
 CC -!- FUNCTION: REQUIRED FOR CHROMOSOMAL REPLICATION.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC CDC7 SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use.

RESULT 6
 DB 77 SPDP -EVETGCGEDIMVRIF-----FAPEAHEKALTGWHHLKHPYGRMEEMKASG 130
 QY 127 EIQLAEQVLLNLYETTDKHLSP-----DGQYVPRIMFVDPSTVRADITGRYSNRLYA 182
 DB 131 LVESVQVEEIVNEPEYTYKLLSQNPIDGHH-----GLAVESEDPHFPSQOLEQ 180
 QY 183 YEPADTALLDNKMKALKLKTTEL 206
 DB 181 DEADKLDFAIQEVKVTIEMKQOV 204
 RESULT 7
 HSK1_SCHPO
 ID HSK1_SCHPO STANDARD: PRT: 507 AA.
 AC P50582; 094678;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cell cycle protein kinase hsk1 (EC 2.7.1.-)
 GN HSK1 OR SPBC776.12C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 CX NCBI_TaxID=4896;
 [1]
 RN SEQUENCE FROM N.A. (LONG ISOFORM).
 RP STRAIN=JY2;
 RC MEDLINE=95347336; PubMed=7621824;
 RA Masai H., Miyake T., Arai K.-I.;
 RT "hsk1+, a Schizosaccharomycetes pombe gene related to Saccharomyces
 RL cerevisiae CDC7, is required for chromosomal replication.";
 RN EMBO J. 14:3094-3104(1995).
 [2]
 RN SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RP STRAIN=972;
 RC MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jageis K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford J., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Benito J.,
 RA Lucas M., Rochet M., Gallard C., Talleda V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RL "The genome sequence of Schizosaccharomycetes pombe.";
 CC Nature 415:871-880(2002).
 CC -!- FUNCTION: REQUIRED FOR CHROMOSOMAL REPLICATION.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC CDC7 SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use.

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL:	D50493;	BAA09087.1;	-	
EMBL:	AL035263;	CAA22885.1;	-	
EMBL:	AL035263;	CAB38417.1;	-	
InterPro:	IPR000719;	Euk_pkinase.		
InterPro:	IPR002290;	Ser_thr_pkinase.		
Pfam:	PF00069;	pkinaase;	1.	
ProDom:	PD000001;	Euk_pkinase;	1.	
PROSITE:	P500107;	PROTEIN_KINASE_ATP;	FALSE_NEG.	
PROSITE:	P500108;	PROTEIN_KINASE_ST.	1.	
PROSITE:	P50011;	PROTEIN_KINASE_DOM;	1.	
Cell cycle:	Cell division;	Serine/threonine-protein kinase;		
Transferase:	ATP-binding;	Alternative splicing.		
DOMAIN	68	433	PROTEIN KINASE.	
NP_BIND	74	82	ATP (BY SIMILARITY).	
BINDING	129	129	ATP (BY SIMILARITY).	
ACT_SITE	216	216	BY SIMILARITY.	
VARSPLIC	92	119	MISSING (IN SHORT ISOFORM).	
SEQUENCE	507 AA.	58407 MD.	04970558218441EE.CBCEA.	

Query Match		7.5%; Score 80; DB 1; Length 843;
Best Local Similarity		21.3%; Pred. No. 14;
Matches		38; Conservative 29; Mismatches 57; Indels 54; Gaps 7;
QY	20	TQGRWRKSRVAMEKIPVSFLLIVA-----LSYTLARDITVKPGAKKDKTDS 67
DB	453	TANRHQTEKTSQEKIKAS-FDALVARMOTKVAERETLLEDTKSKETFEKKOMQOQLKDE 511
QY	68	RPKLPTLSRGWDQLTWTQTYEALYKSKTSNKPMLIIHLDCEPHSQALKKVF AE--- 124
DB	512	KARLDQDLEE-----WGKCEQDITEARKEQELPKPYH-DL ANAEAEHKTILVE RD 563
QY	125	--NKEIQKLAEQFV-----LLNL--VYETTDKHLSPD 152
DB	564	EINAEISRLQDAIVDHKRKISGYGNLDLDAQNRNIREDDKLELGQTRESLESHLND 621
RESULT 12		
ERA_HELPJ		
ID	ERA_HELPJ	STANDARD; PRT; 301 AA.
AC	O92LW0;	
DT	16-OCT-2001 (Rel. 40, Created)	
DT	16-OCT-2001 (Rel. 40, Last sequence update)	
DT	16-OCT-2001 (Rel. 40, Last annotation update)	
DE	GTP-binding protein era homolog.	
GN	ERA OR JHP0466.	
OS	Helicobacter pylori J99 (Campylobacter pylori J99).	
OC	Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;	
OC	Helicobacter.	
OX	NCBI_TaxID=85963;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=99120557; PubMed=9223682;	
RA	Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,	
RA	Smith D.R., Noonan B., Guld B.C., deJonge B.L., Carmel G.,	
RA	Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,	
RA	Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,	
RA	Trust T.J.;	
RT	"Genomic sequence comparison of two unrelated isolates of the human	
RT	gastric pathogen Helicobacter pylori.";	
RL	Nature 37:176-180(1999).	
CC	-1- FUNCTION: BINDS BOTH GDP AND GTP, HAS AN INTRINSIC GTPASE ACTIVITY	
CC	AND IS ESSENTIAL FOR CELL GROWTH (BY SIMILARITY).	
CC	-1- SIMILARITY: BELONGS TO THE ERA/TRME FAMILY OF GTP-BINDING	
CC	PROTEINS. ERA SUBFAMILY.	
CC	-1- SIMILARITY: CONTAINS 1 KH TYPE-2 DOMAIN.	
CC		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC		
DR	EMBL; AE001480; AAD06035.1; -	
DR	HSP; P06616; 1EGA.	
DR	InterPro; IPR004044; KH_TYPE_2.	
DR	InterPro; IPR004087; KH_dom.	
DR	InterPro; IPR005225; Small_GTP.	
DR	Pfam; PF00013; KH-domain; 1.	
DR	TIGRFAMs; TIGR00231; small_GTP; 1.	
DR	TIGRFAMs; TIGR00436; era; 1.	
DR	PROSITE; PS50823; KH_TYPE_2; 1.	
KW	GTP-binding; RNA-binding; Complete proteome.	
FT	NP_BIND 12 19 GTP (POTENTIAL).	
FT	NP_BIND 64 68 GTP (POTENTIAL).	
FT	NP_BIND 122 125 GTP (POTENTIAL).	
FT	DOMAIN 204 280 KH TYPE-2.	
SQ	SEQUENCE 301 AA; 34485 MW; F668C7730882EDBF CRC64;	
Query Match,		7.5%; Score 79.5; DB 1; Length 301;

Best Local Similarity		27.8%; Pred. No. 4.2;
Matches		25; Conservative 13; Mismatches 33; Indels 19; Gaps 4;
QY	89	YEEALYKSKTSNKP-LMITHLDECPHSQALKKVF AEKIQKLAEQFVL----- 137
DB	106	YEEFL---SLCQKPHILALS KIDTATHKQVLQKL---QEYQKYSOFTLALVPLSAKKSQ 158
QY	138	-LNLVYETTDKHLSPDGQVYPRIMEFVDP SL 166
DB	159	NLNALLECISKHLSPSAWLFEKDLMSDEKM 188
RESULT 13		
CBP1_ORYSA		
ID	CBP1_ORYSA	STANDARD; PRT; 510 AA.
AC	P37890;	
DT	01-OCT-1994 (Rel. 30, Created)	
DT	01-OCT-1994 (Rel. 30, Last sequence update)	
DT	15-JUN-2002 (Rel. 41, Last annotation update)	
DE	Serine carboxypeptidase I precursor (EC 3.4.16.5) (Carboxypeptidase	
DE	C).	
GN	CBP1.	
OS	Oryza sativa (Rice).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
OC	Ehrhartoideae; Oryzaceae; Oryza.	
OX	NCBI_TaxID=4530;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	STRAIN=cv. Yukihikari;	
RX	MEDLINE=94213891; PubMed=8161571;	
RA	Washio K., Ishikawa K.;	
RT	"Cloning and sequencing of the gene for type I carboxypeptidase in	
RT	rice.";	
RL	Biochim. Biophys. Acta 1199:311-314(1994).	
CC	-1- CATALYTIC ACTIVITY: Release of a C-terminal amino acid with a	
CC	broad specificity.	
CC	-1- PTM: THREE DISULFIDE BONDS ARE PRESENT (POTENTIAL).	
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S10.	
CC		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC		
DR	EMBL; D17586; BAA04510.1; -	
DR	PIR; S43516; S43516.	
DR	HSP; P08819; 1WHT.	
DR	MEROPS; S10.004; -	
DR	InterPro; IPR000379; Ser_estrs_site.	
DR	InterPro; IPR001563; Serine_carbpept.	
DR	Pfam; PF00450; serine_carbpept; 1.	
DR	PRINTS; PR00724; CRBOXYPTASEC.	
DR	ProDom; PD001189; Serine_carbpept; 2.	
DR	PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.	
DR	PROSITE; PS00342; MICROBODIES_CTER; 1.	
DR	PROSITE; PS00560; CARBOXYPEPT_SER_HIS; 1.	
KW	Hydrolase; Carboxypeptidase; Glycoprotein; zymogen; Signal.	
FT	SIGNAL 1 25 POTENTIAL.	
FT	PROPEP 26 36 POTENTIAL.	
FT	CHAIN 37 510 SERINE CARBOXYPEPTIDASE I.	
FT	ACT_SITE 194 194 BY SIMILARITY.	
FT	ACT_SITE 434 434 BY SIMILARITY.	
FT	ACT_SITE 487 487 BY SIMILARITY.	
FT	CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	SITE 508 510 MICROBODY TARGETING SIGNAL (POTENTIAL).	
SQ	SEQUENCE 510 AA; 55709 MW; 28896247FA1371CF CRC64;	

```

Query Match          7.4%; Score 79; DB 1; Length 510;
Best Local Similarity 37.5%; Pred. No. 9.1;
Matches 30; Conservative 5; Mismatches 33; Indels 12; Gaps 4;

QY 73 QTLRG--WGDLIWTQTYEALYKSTSNKPLMIHHLDECHPSQALKKVAENKEIOK 130
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DQ 259 QTACHGYNWNT---TDKCNALYKVDTSINDLNIDYDLEPCYHSKTIKVTTPAN---TK 312
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 131 LAEQFVLINLVYETTDKHL 150
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DQ 313 LPKSFQHLG-----TTTKPLA 328
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
AROE_BACSU          STANDARD; PRT: 280 AA.
ID AROE_BACSU
AC P54374;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Shikimate 5-dehydrogenase (EC 1.1.1.25).
GN AROD.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=97124195; PubMed=8969508;
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M., Kobayashi Y.
RA "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of the Bacillus subtilis genome containing the skin element and many sporulation genes."
RL Microbiology 142:3103-3111(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschii C.V., Caldwell B., Capuano N.J., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel N.A., Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T., Entian K.D., Errington J., Fabret C., Ferrarini E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C., Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B., Sorokin A., Taccani E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RA "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis."
RL Nature 390:249-256(1997).
CC -!- CATALYTIC ACTIVITY: Shikimate + NADP(+) = 5-dehydroshikimate + NADPH.
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;

```

```

fourth step.
-!- SIMILARITY: BELONGS TO THE SHIKIMATE DEHYDROGENASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
EMBL: D84432; BAA12445.1; -
DR EMBL: Z99117; CAB14508.1; -
DR Subtilist; BG11522; arod.
DR InterPro; IPR002907; Shikimate_DH.
DR Pfam; PF01488; Shikimate_DH; 1.
DR TIGRFAMs; TIGR00507; arod; 1.
KW Aromatic amino acid biosynthesis; Oxidoreductase; NADP;
SQ COMPLETE PROTEOME.
SEQUENCE 280 AA; 30642 MW; 1EE0CE9C2EDA553F CRC64;

Query Match          7.4%; Score 78.5; DB 1; Length 280;
Best Local Similarity 24.8%; Pred. No. 4.7;
Matches 39; Conservative 17; Mismatches 58; Indels 43; Gaps 5;

QY 80 GDQLIWTQTYEALYKS--KTSNKLPLMIHHL-----DECPH 114
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DQ 93 GDLKGYNTDGEFVSLMKVLDKPISELSFLMIGAGGAARAIFTTFRVNTPKKPDICNR 152
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 115 SQALKKVFAP-----NKE---IQKLAQFVLLNLVYETTDKHLSPGQGVPRIMFVDP 165
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DQ 153 TLEKARLLEATPSFHKNELVSIKEAEERLEQYDIHTTSVGMYPNVDDVPLSLQRAAS 212
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 166 LIVRADITGRYSNRLYAYEPADTALLDNMKALKLL 202
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DQ 213 SAVVCDI-----VYNPIQTALLKEASQKGLKTL 240
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
TPP2_MOUSE
ID TPP2_MOUSE          STANDARD; PRT: 1262 AA.
AC O64514;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tripeptidyl-peptidase II (EC 3.4.14.10) (TPP-II) (Tripeptidyl amino-peptidase).
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Leaden X A1;
RX MEDLINE=95091686; PubMed=7998988;
RA Tomkinson B.;
RT "Characterization of cDNA for murine tripeptidyl-peptidase II reveals alternative splicing."
RL Biochem. J. 304:517-523(1994).
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal tripeptide from a polypeptide.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to licensee@sib-sib.ch).
CC -----
DR EMBL; X81323; CAA57103.1; -.
DR HSSP; P00782; 2SBT.
DR MEROPS; S08.090; -.
DR MGD; MGI:102724; Tpp2.
DR InterPro; IPR000209; Peptidase_s8.
DR Pfam; PF00082; Peptidase_s8; 3.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; FALSE_NEG.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Alternative splicing.
FT ACT_SITE 44 44 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 264 264 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 449 449 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT VARSPLIC 985 997 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 1262 AA; 139878 MW; D50D22C85544B034 CRC64;

Query Match 7.4%; Score 78.5; DB 1; Length 1262;
Best Local Similarity 22.9%; Pred. No. 32;
Matches 44; Conservative 31; Mismatches 78; Indels 39; Gaps 7;

QY 5 ACTLGGSGPHPSRLTQGRVVRKSRVAMEKIPVSFLLLVALLSYTLARDTTPKPGAK--- 61
Db 970 AGSLTLKTELKKGKAGQSAKROGKFKKDVIPVHYL-----IPPTKIRNGSKDKE 1021

QY 62 -----KDTKDSRPKLPQTLSRGWGDLQIWTQYEEALYKSKTSNKLMI--IHHLDECP 113
Db 1022 KDSEKEDLKEEFTEALRDLKIOWMTKLDSTDIYNE-LKETYPAYLPLYVARLHQLD--- 1077

QY 114 HSQALKKVFPAENKEIQKLAQFVLLNLVYETTDKHLSPDQYVPRIMFVDP---SLTVRA 170
Db 1078 -----AERKMKRLNEIVDAANAVI-----SHIDQALAVYIAMKTDPRPDAAATIKN 1124

QY 171 DITGRYSNRLYA 182
Db 1125 DMDKQKSLIDA 1136

```

Search completed: January 17, 2003, 12:57:42
Job time : 14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 12:55:16 : Search time 30 seconds

(without alignments)
1414.857 Million cell updates/sec

Title: US-09-646-778a-288

Perfect score: 1063

Sequence: 1 RLSGAGTLSSGSPHPSRLT.....DTALLDNMKKALKLKTTEL 206

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhch:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	899	84.6	175	4	095994	095994 homo sapien
2	828	77.9	175	11	088312	088312 mus musculu
3	555	52.2	165	11	08R3W7	08R3W7 mus musculu
4	555	52.2	166	4	08TD06	08td06 homo sapien
5	548	51.6	159	13	090Y05	090y05 xenopus lae
6	447	42.1	185	13	042251	042251 xenopus lae
7	227	21.4	172	4	096H50	096h50 homo sapien
8	224.5	21.1	170	11	09CQ00	09cqu0 mus musculu
9	222	20.9	172	4	095881	095881 homo sapien
10	199	18.7	186	5	09NA78	09na78 caenorhabdi
11	165	15.5	257	5	09N5S7	09n5s7 caenorhabdi
12	119	11.2	150	17	027777	027777 methanobact
13	89	8.4	950	11	092455	0924s5 rattus norv
14	88.5	8.3	737	2	09AF04	09af04 frankia sp.
15	87	8.2	348	16	0926X3	0926x3 chlamydia p
16	84.5	7.9	378	16	08XIU3	08xiu3 clostridium

ALIGNMENTS

RESULT 1

095994 ID 095994 PRELIMINARY; PRT; 175 AA.
AC 095994;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Secreted CEMENT gland protein XAG-2 homolog (Anterior GRADIENT 2
DE (XENEPUS LAEVIS) homolog), (XAG-2 homolog long protein).
GN HAG-2/R OR HAG-2/C OR HPC8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BREAST;
RX MEDLINE=99009231; PubMed=9790916;
RA Thompson D.A., Weigel R.J.;
RT "HAG-2, the human homologue of the Xenopus laevis cement gland gene
RT XAG-2, is coexpressed with estrogen receptor in breast cancer cell
RT lines.";
RL Biochem. Biophys. Res. Commun. 251:111-116(1998). 10-1998

17 84.5 7.9 1151 16 0929J2
18 84 7.9 435 5 09BJK8
19 83.5 7.9 918 10 09CA88
20 83.5 7.9 1011 5 09U9P6
21 83.5 7.9 1409 5 09V3A7
22 83.5 7.9 1409 5 09U6I2
23 83 7.8 786 4 09H747
24 83 7.8 976 5 09U158
25 82.5 7.8 788 16 067557
26 81.5 7.7 649 10 09C5B1
27 81.5 7.7 1158 3 09UTR5
28 81.5 7.7 1721 5 08SSQ0
29 81 7.6 601 5 095998
30 81 7.6 610 5 09XVM2
31 80.5 7.6 802 4 08TB22
32 80.5 7.6 875 12 09QG69
33 80 7.5 154 16 09PCE7
34 80 7.5 259 10 09M9H0
35 80 7.5 623 10 09FFR8
36 80 7.5 861 5 09NXX3
37 80 7.5 1051 17 08TMS3
38 79.5 7.5 277 3 013703
39 79.5 7.5 536 10 09SKT6
40 79.5 7.5 642 16 09KFD1
41 79 7.4 275 2 087445
42 79 7.4 337 5 095NH3
43 79 7.4 477 3 013745
44 79 7.4 480 5 09VMD8
45 79 7.4 710 5 08STF7

0929j2 listeria in
09bjk8 caenorhabdi
09ca88 arabidopsis
09u9p6 drosophila
09v3a7 drosophila
09u6i2 drosophila
09h747 homo sapien
09u158 leishmania
067557 aquifex aeo
09c5b1 arabidopsis
09utr5 schizosacch
08ssq0 dictyostell
095998 caenorhabdi
09xvm2 caenorhabdi
08tb22 homo sapien
09qg69 apple mosai
09pce7 xylella fas
09m9h0 arabidopsis
09ffr8 arabidopsis
09nxx3 bombyx mori
08tms3 methanosarc
013703 schizosacch
09skt6 arabidopsis
09kfd1 bacillus ha
087445 brachyspira
095nh3 drosophila
013745 schizosacch
09vmd8 drosophila
08stf7 drosophila

```
DR EMBL; AF007791; AAC77358.1; -
DR EMBL; AF08867; AAF22484.1; -
DR EMBL; BC015503; AAH15503.1; -
DR EMBL; AF115926; AAL54870.1; -
SQ SEQUENCE 175 AA; 19979 MW; F271B1BD377BEE11 CRC64;

Query Match      84.6%; Score 899; DB 4; Length 175;
Best Local Similarity 100.0%; Pred. No. 2.6e-78;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 MEKIPVSALLLVALLVALSYTLARDTTVKPGAKKDTKDSRPKLPQTLSRGWDQLIWTQTYEE 91
Db 1 MEKIPVSALLLVALLVALSYTLARDTTVKPGAKKDTKDSRPKLPQTLSRGWDQLIWTQTYEE 60

Qy 92 ALYKSKTSNKPLMIITHLDECPHSQALKKVFPAENKEIQKLAEOFVLLNLVYETTDKHLSP 151
Db 61 ALYKSKTSNKPLMIITHLDECPHSQALKKVFPAENKEIQKLAEOFVLLNLVYETTDKHLSP 120

Qy 152 DGQYVPRIMFVDPSTVRADITGRYSNRLYAYEPADTALLDNMKKALKLLKTEL 206
Db 121 DGQYVPRIMFVDPSTVRADITGRYSNRLYAYEPADTALLDNMKKALKLLKTEL 175

RESULT 2
O88312 ID O88312 PRELIMINARY; PRT; 175 AA.
AC O88312;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE GOB-4 protein (Anterior GRADIENT 2) (XENOPUS LAEVIS) (XENOPUS LAEVIS).
DE LAEVIS).
GN AGR2 OR GOB-4 OR MAG-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN STRAIN=BA1B/C; TISSUE=INTESTINE;
RP Komiya T., Hirohashi S.;
RA "Cloning of a gene, gob-4, which is expressed in intestinal goblet cells in mice.";
RT Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS WEBSTER;
RX MEDLINE=99009231; PubMed=9790916;
RT Thompson D.A., Weigel R.J.;
RT "hAG-2, the human homologue of the xenopus laevis cement gland gene XAG-2, is coexpressed with estrogen receptor in breast cancer cell lines.";
RL Biochem. Biophys. Res. Commun. 251:111-116(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gliss C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
```

```
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Anterior gradient protein 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Adam P.J., Boyd R., Fletcher G.C., Tyson K., Terrett J.;
RT "Proteomic Characterization of Breast Cancer Cell Membranes.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY069977; AAL55402.1; -
SQ SEQUENCE 166 AA; 19171 MW; BIA4804B47705D65 CRC64;

Query Match 52.2%; Score 555; DB 4; Length 166;
Best Local Similarity 64.7%; Pred. No. 2.3e-45;
Matches 108; Conservative 28; Mismatches 19; Indels 12; Gaps 4;

QY 41 LLLVALSYTLARDTTVKPGAKDKTSRPLQPTLSRGWGDLQIWTQYEEALYKSTSN 100
DB 11 LLLTVSSNLA--IAIK-----KEKRP--PQTLSRGWGDDITWQYEEGLFYAKQSK 59
QY 101 KPLMTIHLDECPSOALKKVFPAENKEIQKLA-EQVLLNLVYETDCKHLSPDGQVVPRI 159
DB 60 KPLMVIHLEDCQYSOALKKVFPAENKEIQKLA-EQVLLNLVYETDCKHLSPDGQVVPRI 119
QY 160 MFVDPSTLVTRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 206
DB 120 MFVDPSTLVTRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 166

RESULT 5
Q90Y05 PRELIMINARY; PRT; 159 AA.
AC Q90Y05;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Cement gland-specific protein CGS.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A.
RA Alexandrova E.M., Novoselov V.V., Zaraisky A.G.;
RT "Three novel genes expressed in the anterior part of the Xenopus
laevis embryo.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF314056; AAL26844.1; -
SQ SEQUENCE 159 AA; 18275 MW; 738D6228483888EB CRC64;

Query Match 51.6%; Score 548; DB 13; Length 159;
Best Local Similarity 60.0%; Pred. No. 1e-44;
Matches 105; Conservative 26; Mismatches 28; Indels 16; Gaps 1;

QY 32 MEKIPVSFLLLVALSYTLARDTTVKPGAKDKTSRPLQPTLSRGWGDLQIWTQYEE 91
DB 1 METVLKSLFLLVATSFYLAKE-----RKPTLSRGWGDNLEWVQYEE 44
QY 92 ALYKSTSNKPLMTIHLDECPSOALKKVFPAENKEIQKLA-EQVLLNLVYETDCKHLS 151
DB 45 GLFRAKSENKPLLNHRNDCPSOALKKAFABHOGIQKLAEEFILLNLVYDPTDKNLQ 104
QY 152 DGQYVPRIMEFVDPSTLVTRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 206
DB 105 DGQYVPRIMEFVDPSTLVTRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 159

RESULT 6
```

```
042251 PRELIMINARY; PRT; 185 AA.
AC 042251;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
DE Cement gland-specific.
GN XAG-2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEAD;
RX MEDLINE=89324051; PubMed=2752418;
RA Sive H.L., Hattori K., Weintraub H.;
RT "Progressive determination during formation of the anteroposterior
axis in Xenopus laevis.";
RL Cell 58:171-180(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=HEAD;
RA Sive H.L., Wainstock D.H., Kennedy B.S.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF025474; AAB81968.1; -
SQ SEQUENCE 185 AA; 20498 MW; D6CDE02DEC3B857B CRC64;

Query Match 42.1%; Score 447; DB 13; Length 185;
Best Local Similarity 50.6%; Pred. No. 6e-35;
Matches 89; Conservative 26; Mismatches 49; Indels 12; Gaps 2;

QY 43 LVALSYTLARDTTVKP-----GAKDKTSRPLQPTLSRGWGDLQIWTQYEE 91
DB 10 LVLLCSVLGEAALRKPKROAGATDTNGAKSEPAVKTKGLTKLDRGWGEDIEWAQYEE 69
QY 92 ALYKSTSNKPLMTIHLDECPSOALKKVFPAENKEIQKLA-EQVLLNLVYETDCKHLS 150
DB 70 GLAKARENKPLMVIHLEDCPYSLALKAFVADMAQKLAQEDFIMLNLVHPVADEKQS 129
QY 151 PDGQYVPRIMEFVDPSTLVTRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 206
DB 130 PDGQYVPRIMEFVDPSTLVTRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 185

RESULT 7
Q96H50 PRELIMINARY; PRT; 172 AA.
AC Q96H50;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Hypothetical 19.2 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC008913; AAB08913.1; -
DR InterPro: IPR000063; ThioRed.
DR PROSITE: PS00194; THIOREDOXIN; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 172 AA; 19228 MW; 5C9DDA6D5A7C4BDF CRC64;

Query Match 21.4%; Score 227; DB 4; Length 172;
Best Local Similarity 34.8%; Pred. No. 6.6e-14;
Matches 55; Conservative 25; Mismatches 56; Indels 22; Gaps 3;
```


QY 66 DSRPKLPOT-----LSRGWQDLIWTQTYEALYKSKTSNKPLMII 106
 Db 2 ETRPRLGATLGLFSFLLVSSDGHNLGKFGFDHIHW-RTLEDGKKEAASGLPLMVI 60

QY 107 HHLDECPHSQALKKVAENKEIOKLABQFVLLNLVYETTDKH--LSPDQGVPRIMFVDP 164
 Db 61 IHKSWGACKALKPKFAESTEISELSHNFVMVNLDEEPEKDEFSPDGGYIPRILFLDP 120

QY 165 SLTVRADITGRYSNRLYAYEPADTALLDNKKALKLL 202
 Db 121 SGKVHPEIINENGNSPSKYFYVSAEQVVQGMKEAQLRL 158

RESULT 8
 Q9CQ00 PRELIMINARY; PRT; 170 AA.
 AC Q9CQ00;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DE 0610040B21Rik protein (RIKEN cDNA 0610040B21 gene).
 GN 0610040B21Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO, AND KIDNEY;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK003481; BAB22811.1; -;
 DR EMBL; AK002862; BAB22413.1; -;
 DR EMBL; BC068857; AAH06857.1; -;
 DR MGP; MGI:1913323; 0610040B21Rik.
 DR InterPro; IPR000886; ER target.
 DR InterPro; IPR000063; Thiores.
 DR PROSITE; PS00014; ER TARGET; UNKNOWN_1.
 DR PROSITE; PS00194; THIOREDOXIN; UNKNOWN_1.
 SQ SEQUENCE 170 AA; 19048 MW; 5B91FC9BE12C5E44 CRC64;

Query Match 21.1%; Score 224.5; DB 11; Length 170;
 Best Local Similarity 34.9%; Pred. No. 1,le-13;
 Matches 58; Conservative 25; Mismatches 58; Indels 25; Gaps 4;

QY 39 AFLLLVALSYTLARDTTVPKAKKDKDSRPKLPQTLPSRGWQDLIWTQTYEALYKSKT 98
 Db 14 SFLLLITSS-----DGR-----TGLGKFGFDHIHW-RTLEDGKKEAAA 50

QY 99 SNKPLNIHHLDECPHSQALKKVAENKEIOKLABQFVLLNL--VYETTDKHLSPDQGV 156

Db 51 SGLPLMVIHHSWCGACKALKPKFAESTEISELSHNFVMVNLDEEPEKDEFSPDGGYI 110
 QY 157 PRIMEVDPSLTVRADITGRYSNRLYAYEPADTALLDNKKALKLL 202
 Db 111 PRILFLDPGKVRPEIINENGNSPSKYFYVSAEQVVQGMKEAQLRL 156

RESULT 9
 O95881 PRELIMINARY; PRT; 172 AA.
 AC O95881;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Unknown (Hypothetical 19.2 kDa protein).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96207227; PubMed=8619474;
 RA Andersson B., Wentland M.A., Ricafrente J.Y., Liu W., Gibbs R.A.;
 RT "A 'double adaptor' method for improved shotgun library
 RT construction."
 RL Anal. Biochem. 236:107-113(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97264341; PubMed=9110174;
 RA Yu W., Andersson B., Worley K.C., Muzny D.M., Ding Y., Liu W.,
 RA Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;
 RT "Large-scale concatenation cDNA sequencing."
 RL Genome Res. 7:353-358(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Mei G., Yu W., Gibbs R.A.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RA Strausberg R.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=COLON;
 RA Strausberg R.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF131738; AAD20035.1; -;
 DR EMBL; BC001493; AAH01493.1; -;
 DR EMBL; BC008953; AAH08953.1; -;
 DR InterPro; IPR000063; Thiores.
 DR PROSITE; PS00194; THIOREDOXIN; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 172 AA; 19206 MW; 3092E9515A7C4094 CRC64;

Query Match 20.9%; Score 222; DB 4; Length 172;
 Best Local Similarity 34.2%; Pred. No. 2e-13;
 Matches 54; Conservative 26; Mismatches 56; Indels 22; Gaps 3;

QY 66 DSRPKLPOT-----LSRGWQDLIWTQTYEALYKSKTSNKPLMII 106
 Db 2 ETRPRLGATLGLFSFLLVSSDGHNLGKFGFDHIHW-RTLEDGKKEAASGLPLMVI 60

QY 107 HHLDECPHSQALKKVAENKEIOKLABQFVLLNL--VYETTDKHLSPDQGVPRIMFVDP 164
 Db 61 IHKSWGACKALKPKFAESTEISELSHNFVMVNLDEEPEKDEFSPDGGYIPRILFLDP 120

QY 165 SLTVRADITGRYSNRLYAYEPADTALLDNKKALKLL 202
 Db 121 SGKVHPEIINENGNSPSKYFYVSAEQVVQGMKEAQLRL 158

```

RESULT 10
Q9NA78 PRELIMINARY; PRT; 186 AA.
AC Q9NA78; 2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Y57A10A.23 protein.
GN Y57A10A.23
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Smye R.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AL17195; CAB55026.1; -.
DR InterPro; IPR000063; Thiorod.
SQ SEQUENCE 186 AA; 21467 MW; F94E13D2F773321E CRC64;

Query Match 18.7%; Score 199; DB 5; Length 186;
Best Local Similarity 30.0%; Pred. No. 3.5e-11;
Matches 51; Conservative 38; Mismatches 57; Indels 24; Gaps 6;

QY 41 LLLVALSYTLARDTTVKGAKKDTKDSRPKLPTQLSRGWGDLIWTQTYEALYKSTSN 100
Db 4 LLLALL-----VSASAYAFDKIKDSIONPLRGFGDDIAWK-WEDAIETALD 53
QY 101 KPLMIHHLDECPHSAQALKVFAEN---KEIOKLAQFVLLNLV--YETTDKHLSPDGOY 155
Db 54 KPFIILLHKSCHACKALKKTKQSNAKAFKLSHFVVMVNTEDDEPPEEYRPGKY 113
QY 156 VPRIMEVDPSTVRADITGRYSNRL-----YAYEPADTALLDNMKKAL 200
Db 114 IPRLLFLDKN-----GDLLQEFKNKAENYAYYSSPADILNSMKDVLK 159

RESULT 11
Q9NS57 PRELIMINARY; PRT; 257 AA.
AC Q9NS57;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE Hypothetical 30.2 kDa protein.
GN F49H12.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Woldmann P., Nhan M., Chissoe S., Hawkins J.;
RA "The sequence of C. elegans cosmid F49H12.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL EMBL; AC006642; AAF39830.1; -.
DR InterPro; IPR000886; ER_target.
DR InterPro; IPR000063; Thiorod.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 257 AA; 30205 MW; BF7FA62D04DBED6 CRC64;

Query Match 15.5%; Score 165; DB 5; Length 257;
Best Local Similarity 30.7%; Pred. No. 9.8e-08;
Matches 51; Conservative 35; Mismatches 58; Indels 22; Gaps 8;

QY 41 LLLVALSYTLARDTTVKGAKKDTKDSRPKLPTQLSRGWGDLIWTQTYEALYKSTSN 100
Db 3 LLLLSL---LAVAVLAKKESKEIKD-----LSHGYPASIEWVE-FDKAIGIAKDLN 50
QY 101 KPLMIHHLDECPHSAQALK---KVFAENKEIOKLAQFVLLNLV--YETTDKHLSPDGOY 155
Db 51 KPFIILLHKSCHACKALKKRELKSSPKTDELILSRKPMVNVDEDEPDPKSPDGOY 110
QY 156 VPRIMEVDPSTVRADITGRYSNRLYAYE-PADTALLDNMKKAL 199
Db 111 IPRLLFLDKN-----GDLLQEFKNKAENYAYYSSPADILNSMKDVLK 153

RESULT 12
Q27777 PRELIMINARY; PRT; 150 AA.
AC Q27777;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Protein DISULPHIDE isomerase.
GN MTH1745.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
DR EMBL; AF000930; AAB86215.1; -.
DR InterPro; IPR000063; Thiorod.
DR Pfam; PF00085; Thiorod; 1.
KW PROSITE; PS00194; THIOREDOXIN;
KW Isomerase; Complete proteome.
SQ SEQUENCE 150 AA; 16939 MW; E0672E64B7E340EB CRC64;

Query Match 11.2%; Score 119; DB 17; Length 150;
Best Local Similarity 27.1%; Pred. No. 0.0012;
Matches 35; Conservative 28; Mismatches 46; Indels 20; Gaps 4;

QY 39 AFLLVALSYYTLARDTTVKGAKKDTKDSRPKLPTQLSRGWGDLIWTQTYEALYKSKT 98
Db 19 AGLSLVLIIVTV-----QPRVPSLTDEKD-LKWTYEHDEAIKEASR 60
QY 99 SNKPLMIHHLDECPHSAQALKVFAENKEIQ-KLAQFVLLNLVYETTDKHLSPDGOY-V 156

```

Query Match 8.2%; Score 87; DB 16; Length 348;

	Best Local Similarity	25.0%;	Pred. No. 4.4;	Indels	24;	Gaps	4;
	Matches	29;	Conservative	26;	Mismatches	37;	
QY	29	RVAMEKIPVS	AFILLVALSYTLARDTTVKQCAKDKYKDSRPKLPTQLSRGWGDLIIWTQT	88			
Db	9	RCSLKQLKVLATLL	--LSLSL--PTLEAENRDS-----DSIVWHLD	46			
QY	89	YERALKYSKTSNPKLMIHHLDEC--PHSQALKKVFPAENKEIOKLARQFVLLNVY	142				
Db	47	YQBALQSKRAEPLAVIFSGSDWNGPCMKIRREVLESPEFIRKQGVFCVCEVEY	102				

Search completed: January 17, 2003, 12:56:34
Job time : 32 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 12:57:01 ; Search time 11 seconds
(without alignments)
372.270 Million cell updates/sec

Title: US-09-646-778A-288
Perfect score: 1063
Sequence: 1 RLSAGTLGSGPHPSRRLLT.....DTALLDNMKNKALKLKTTEL 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 120991 seqs, 19878514 residues
Total number of hits satisfying chosen parameters: 120991

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pap.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pap.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pap.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pap.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pap.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap.*
7: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pap.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pap.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pap.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pap.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	899	84.6	175	10	US-09-792-439-2
2	555	52.2	166	9	US-10-066-500-54
3	555	52.2	166	9	US-09-956-751-1
4	555	52.2	168	9	US-10-025-380-631
5	555	52.2	168	10	US-09-922-217-631
6	555	52.2	168	10	US-09-833-263-631
7	222	20.9	172	10	US-09-954-846-1
8	81.5	7.7	901	10	US-09-892-287-5
9	81.5	7.7	968	10	US-09-969-528-7
10	78	7.3	453	9	US-09-981-353-48
11	77	7.2	1049	10	US-09-815-242-10732
12	71.5	6.7	289	10	US-09-815-242-13244
13	71.5	6.7	289	10	US-09-815-242-13582
14	71	6.7	306	9	US-10-017-910-7
15	71	6.7	529	10	US-09-849-566-6
16	71	6.7	529	10	US-09-849-566-8
17	71	6.7	529	10	US-09-907-859-6
18	71	6.7	529	10	US-09-907-859-8
19	70.5	6.6	481	10	US-09-994-185-2

20	70.5	6.6	481	12	US-10-004-139-2	Sequence 2, Appli
21	70.5	6.6	917	9	US-10-001-835-189	Sequence 189, App
22	70.5	6.6	1252	9	US-10-047-542-89	Sequence 89, Appl
23	70	6.6	185	10	US-09-881-752A-40	Sequence 40, Appl
24	70	6.6	1198	12	US-10-001-215-2	Sequence 2, Appli
25	69	6.5	1275	9	US-10-025-201-3	Sequence 3, Appli
26	69	6.5	2905	8	US-08-424-550B-401	Sequence 401, App
27	68.5	6.4	302	10	US-09-815-242-11334	Sequence 11334, A
28	68.5	6.4	456	10	US-09-861-400-12	Sequence 12, Appl
29	68.5	6.4	481	10	US-09-861-400-11	Sequence 11, Appl
30	68.5	6.4	809	9	US-10-105-695-2	Sequence 2, Appli
31	68.5	6.4	809	9	US-10-105-694-2	Sequence 2, Appli
32	68.5	6.4	809	12	US-09-747-521-2	Sequence 2, Appli
33	68.5	6.4	809	12	US-10-106-014-2	Sequence 2, Appli
34	67.5	6.3	481	10	US-09-731-872-271	Sequence 271, App
35	67	6.3	257	10	US-09-815-242-11278	Sequence 11278, A
36	67	6.3	578	9	US-09-912-672A-27	Sequence 27, Appl
37	67	6.3	3092	10	US-09-801-368-172	Sequence 172, App
38	66.5	6.3	270	10	US-09-350-874-14	Sequence 14, Appl
39	66	6.2	453	9	US-09-738-626-4378	Sequence 4378, Ap
40	66	6.2	772	9	US-09-738-626-6661	Sequence 6661, Ap
41	65.5	6.2	305	10	US-09-895-072-3	Sequence 3, Appli
42	65.5	6.2	305	10	US-09-986-552-3	Sequence 3, Appli
43	65.5	6.2	448	10	US-09-815-108-6	Sequence 6, Appli
44	65.5	6.2	472	10	US-09-815-108-5	Sequence 5, Appli
45	65.5	6.2	504	9	US-09-992-598-119	Sequence 119, App

ALIGNMENTS

RESULT 1
US-09-792-439-2
; Sequence 2, Application US/09792439
; Patent No. US20020111303A1
; GENERAL INFORMATION:
; APPLICANT: Boyd, Robert Simon
; APPLICANT: Stamps, Alasdair Craig
; APPLICANT: Terrett, Jonathan Alexander
; APPLICANT: Tyson, Kerry Louise
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
; FILE REFERENCE: 2543-1-005N
; CURRENT APPLICATION NUMBER: US/09/792.439
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: GB 0004576.5
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 175
; TYPE: PRT
; ORGANISM: human
US-09-792-439-2

Query Match	84.6%	Score	899	DB	10	Length	175
Best Local Similarity	100.0%	Pred. NO.	2.7e-86				
Matches	175	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	32	MEKIPVSFAILLVALSYTLARDITVVKGAKKDTKDSRKLPTLSRGWDQLIWTQYEE	91				
Db	1	MEKIPVSFAILLVALSYTLARDITVVKGAKKDTKDSRKLPTLSRGWDQLIWTQYEE	60				
QY	92	ALYKSTSNKPLMTIHHLDPCPHSQALKKVFPAENKEIOKLABQFVLLNLVYETTDKHLSP	151				
Db	61	ALYKSTSNKPLMTIHHLDPCPHSQALKKVFPAENKEIOKLABQFVLLNLVYETTDKHLSP	120				
QY	152	DGQVPRIMEVDFSLTVRADITGRYSNRLYAYEPADTALLDNMKNKALKLKTTEL	206				
Db	121	DGQVPRIMEVDFSLTVRADITGRYSNRLYAYEPADTALLDNMKNKALKLKTTEL	175				

RESULT 2
US-10-066-500-54

```

; Sequence 54, Application US/10066500
; Patent No. US20020177163A1
; GENERAL INFORMATION:
; APPLICANT: Avi J. Ashkenazi
; APPLICANT: Kevin P. Baker
; APPLICANT: David A. Botstein
; APPLICANT: Luc Desnoyers
; APPLICANT: Dan L. Eaton
; APPLICANT: Sherman Fong
; APPLICANT: Wei-Qiang Gao
; APPLICANT: Hanspeter Gerber
; APPLICANT: Mary E. Gerritsen
; APPLICANT: Audrey Goddard
; APPLICANT: Paul J. Godowski
; APPLICANT: Austin L. Gurney
; APPLICANT: Ivar J. Kljavin
; APPLICANT: Jennie P. Mather
; APPLICANT: Mary A. Napier
; APPLICANT: James Pan
; APPLICANT: Nicholas F. Paoni
; APPLICANT: Margaret Ann Roy
; APPLICANT: Timothy A. Stewart
; APPLICANT: Daniel Tumas
; APPLICANT: Colin K. Watanabe
; APPLICANT: P. Mickey Williams
; APPLICANT: William I. Wood
; APPLICANT: Zemin Zang
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3130RIC7
; CURRENT APPLICATION NUMBER: US/10/066,500
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 10/002,796
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066840
; PRIOR FILING DATE: 1997-11-25
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/095998
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/097000
; PRIOR FILING DATE: 1998-08-18
; PRIOR APPLICATION NUMBER: 60/099601
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099811
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100858
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101922
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106032
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/109304
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/125778
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/139695
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/145070
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/149396
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 08/918874
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 08/933821
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 08/960507
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 09/114844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: 09/136801
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: 09/136804
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: 09/136828
; PRIOR FILING DATE: 1998-08-19
; PRIOR APPLICATION NUMBER: 09/158342
; PRIOR FILING DATE: 1998-09-21
; PRIOR APPLICATION NUMBER: 09/180997
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 09/202088
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: 09/254311
; PRIOR FILING DATE: 1999-03-03
; PRIOR APPLICATION NUMBER: 09/254460
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: 09/254465
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: 09/284663
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: 09/332928
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: 09/332929
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: 09/333075
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: 09/333077
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380138
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/380139
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/403296
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 09/403297
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 09/423741

```


; PRIOR FILING DATE: 1999-11-10
 ; PRIOR APPLICATION NUMBER: 09/423844
 ; PRIOR FILING DATE: 1999-11-12
 ; PRIOR APPLICATION NUMBER: 09/522342
 ; PRIOR FILING DATE: 2000-03-09
 ; PRIOR APPLICATION NUMBER: 09/548815
 ; PRIOR FILING DATE: 2000-04-13
 ; PRIOR APPLICATION NUMBER: 09/664610
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: 09/665350
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: 09/709238
 ; PRIOR FILING DATE: 2000-11-08
 ; PRIOR APPLICATION NUMBER: 09/767609
 ; PRIOR FILING DATE: 2001-01-22
 ; PRIOR APPLICATION NUMBER: 09/802706
 ; PRIOR FILING DATE: 2001-03-09
 ; PRIOR APPLICATION NUMBER: 09/808689
 ; PRIOR FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: 09/866028
 ; PRIOR FILING DATE: 2001-05-25
 ; PRIOR APPLICATION NUMBER: 09/870574
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: 09/872035
 ; PRIOR FILING DATE: 2001-06-01
 ; PRIOR APPLICATION NUMBER: 09/886342
 ; PRIOR FILING DATE: 2001-06-19
 ; PRIOR APPLICATION NUMBER: PCT/US98/14552
 ; PRIOR FILING DATE: 1998-07-14
 ; PRIOR APPLICATION NUMBER: PCT/US98/18824
 ; PRIOR FILING DATE: 1998-09-10
 ; PRIOR APPLICATION NUMBER: PCT/US98/19093
 ; PRIOR FILING DATE: 1998-09-14
 ; PRIOR APPLICATION NUMBER: PCT/US98/19330
 ; PRIOR FILING DATE: 1998-09-16
 ; PRIOR APPLICATION NUMBER: PCT/US98/19437
 ; PRIOR FILING DATE: 1998-09-17
 ; PRIOR APPLICATION NUMBER: PCT/US98/24855
 ; PRIOR FILING DATE: 1998-11-20
 ; PRIOR APPLICATION NUMBER: PCT/US98/25108
 ; PRIOR FILING DATE: 1998-12-01
 ; PRIOR APPLICATION NUMBER: PCT/US98/25190
 ; PRIOR FILING DATE: 1998-11-25
 ; PRIOR APPLICATION NUMBER: PCT/US99/05028
 ; PRIOR FILING DATE: 1999-03-08
 ; PRIOR APPLICATION NUMBER: PCT/US99/12252
 ; PRIOR FILING DATE: 1999-06-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/20111
 ; PRIOR FILING DATE: 1999-09-01
 ; PRIOR APPLICATION NUMBER: PCT/US99/20594
 ; PRIOR FILING DATE: 1999-09-08
 ; PRIOR APPLICATION NUMBER: PCT/US99/21090
 ; PRIOR FILING DATE: 1999-09-15
 ; PRIOR APPLICATION NUMBER: PCT/US99/21547

Query Match 52.2%; Score 555; DB 9; Length 166;
 Best Local Similarity 64.7%; Pred. No. 1.6e-50;
 Matches 108; Conservative 28; Mismatches 19; Indels 12; Gaps 4;
 QY 41 LLLVALSYTLARDTTVKPGAKKDTKDSRPKLPQTLSRGWGDDITWVQTYEEGLFYAKSK 100
 Db 11 LLLVTSSNLA--IAIK-----KEKRP--PQTLSRGWGDDITWVQTYEEGLFYAKSK 59
 QY 101 KPLMIHLDCEPHSOALKKVFPAENKEIOKLAE-QFVLLNLVYETTDKHLSPDQGYVPRI 159
 Db 60 KPLMVIHLEDCQYSQALKKVFPAENKEIOEQAQNKFIMLNLMHETTDKNLSLSPDQGYVPRI 119
 QY 160 MFVDPSTLVRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 206
 Db 120 MFVDPSTLVRADITGRYSNRLYAYEPADTALLDNMKKALKLRIQSEL 166
 RESULT 3

US-09-956-751-1
 ; Sequence 1, Application US/09956751
 ; Patent No. US20020055186A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Oxford Glycosciences
 ; TITLE OF INVENTION: Detection of Peptides
 ; FILE REFERENCE: 2754/11022US1
 ; CURRENT APPLICATION NUMBER: US/09/956,751
 ; CURRENT FILING DATE: 2001-09-19
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 166
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-956-751-1

Query Match 52.2%; Score 555; DB 10; Length 166;
 Best Local Similarity 64.7%; Pred. No. 1.6e-50;
 Matches 108; Conservative 28; Mismatches 19; Indels 12; Gaps 4;
 QY 41 LLLVALSYTLARDTTVKPGAKKDTKDSRPKLPQTLSRGWGDDITWVQTYEEGLFYAKSK 100
 Db 11 LLLVTSSNLA--IAIK-----KEKRP--PQTLSRGWGDDITWVQTYEEGLFYAKSK 59
 QY 101 KPLMIHLDCEPHSOALKKVFPAENKEIOKLAE-QFVLLNLVYETTDKHLSPDQGYVPRI 159
 Db 60 KPLMVIHLEDCQYSQALKKVFPAENKEIOEQAQNKFIMLNLMHETTDKNLSLSPDQGYVPRI 119
 QY 160 MFVDPSTLVRADITGRYSNRLYAYEPADTALLDNMKKALKLKTTEL 206
 Db 120 MFVDPSTLVRADITGRYSNRLYAYEPADTALLDNMKKALKLRIQSEL 166

RESULT 4
 US-10-025-380-631
 ; Sequence 631, Application US/10025380
 ; Publication No. US20020182191A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Secrist, Heather
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Meagher, Madeleine Joy
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Jiang, Yuqiu
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Clapper, Jonathan D.
 ; APPLICANT: Skeiky, Yasir A. W.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Vedwick Thomas S.
 ; APPLICANT: Carter, Darrick
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
 ; FILE REFERENCE: 210121.471C14
 ; CURRENT APPLICATION NUMBER: US/10/025,380
 ; CURRENT FILING DATE: 2001-12-19
 ; NUMBER OF SEQ ID NOS: 1129
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 631
 ; LENGTH: 168
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-025-380-631

Query Match 52.2%; Score 555; DB 9; Length 168;
 Best Local Similarity 64.7%; Pred. No. 1.7e-50;
 Matches 108; Conservative 28; Mismatches 19; Indels 12; Gaps 4;
 QY 41 LLLVALSYTLARDTTVKPGAKKDTKDSRPKLPQTLSRGWGDDITWVQTYEEGLFYAKSK 100

```
Db 13 LLLVTSSNLA--IAIK-----KEKRP--PQTLSRGWGDDITWVQTYEGLFYAQSK 61
QY 101 KPLMIHHLEDCPHSQALKKVFQAEKQKLA--QFVLLNLVYETTDKHLSPDQGYVPRI 159
Db 62 KPLMVIHLEDCQYQALKKVFQAEKQKLA--QFVLLNLVYETTDKHLSPDQGYVPRI 121
QY 160 MFVDPSTVTRADITGRYSNRLYAYEPADTALLDNMKKALKLTEL 206
Db 122 MFVDPSTVTRADITGRYSNRLYAYEPADTALLDNMKKALKLTEL 168

RESULT 5
US-09-922-217-631
; Sequence 631, Application US/09922217
; Patent No. US2002007641A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jlangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922.217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 631
; TYPE: PRT
; LENGTH: 168
; ORGANISM: Homo sapiens
US-09-922-217-631

Query Match 52.2%; Score 555; DB 10; Length 168;
Best Local Similarity 64.7%; Pred. No. 1.7e-50;
Matches 108; Conservative 28; Mismatches 19; Indels 12; Gaps 4;

QY 41 LLLVALSYTLARDITVKPGAKDKTQSRPKLPQTLSRGWGDDITWVQTYEGLFYAQSK 100
Db 13 LLLVTSSNLA--IAIK-----KEKRP--PQTLSRGWGDDITWVQTYEGLFYAQSK 61
QY 101 KPLMIHHLEDCPHSQALKKVFQAEKQKLA--QFVLLNLVYETTDKHLSPDQGYVPRI 159
Db 62 KPLMVIHLEDCQYQALKKVFQAEKQKLA--QFVLLNLVYETTDKHLSPDQGYVPRI 121
QY 160 MFVDPSTVTRADITGRYSNRLYAYEPADTALLDNMKKALKLTEL 206
Db 122 MFVDPSTVTRADITGRYSNRLYAYEPADTALLDNMKKALKLTEL 168

RESULT 6
US-09-833-263-631
; Sequence 631, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833.263
; CURRENT FILING DATE: 2001-04-10
```

```
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 631
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-833-263-631

Query Match 52.2%; Score 555; DB 10; Length 168;
Best Local Similarity 64.7%; Pred. No. 1.7e-50;
Matches 108; Conservative 28; Mismatches 19; Indels 12; Gaps 4;

QY 41 LLLVALSYTLARDITVKPGAKDKTQSRPKLPQTLSRGWGDDITWVQTYEGLFYAQSK 100
Db 13 LLLVTSSNLA--IAIK-----KEKRP--PQTLSRGWGDDITWVQTYEGLFYAQSK 61
QY 101 KPLMIHHLEDCPHSQALKKVFQAEKQKLA--QFVLLNLVYETTDKHLSPDQGYVPRI 159
Db 62 KPLMVIHLEDCQYQALKKVFQAEKQKLA--QFVLLNLVYETTDKHLSPDQGYVPRI 121
QY 160 MFVDPSTVTRADITGRYSNRLYAYEPADTALLDNMKKALKLTEL 206
Db 122 MFVDPSTVTRADITGRYSNRLYAYEPADTALLDNMKKALKLTEL 168

RESULT 7
US-09-954-846-1
; Sequence 1, Application US/09954846
; Patent No. US20020102654A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; Corley, Neil C.
; Guegler, Karl J.
; Patterson, Chandra
; Baughn, Mariah R.
; TITLE OF INVENTION: THIOREDOXIN PROTEINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/954,846
; FILING DATE: 17-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/107,248
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0556 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-855-0572
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 172 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT02
; CLONE: 1925679
```

Qy 55 TVKPGAKKDTKDSRPK--LPQTLSRG

Matches 45; Conservative 32;

55 TVKPCAKKDTKDSRPK--I.POTI.SRGWGDOL.TWOTYEAL-YKS-----KTS-NKP

Db 551 TYKDSKTDRWSSCKCRVP-----AWCDRLWRGTNVNQLNRYRSHMLLKTSDHKRPVSAL 605
QY 107 HHK-----DECPHQALKKVAENKEIQKLAEE-----QFVLLNLVYETDK-- 147
Db 606 FHIGVKVDE-----RRYKVEDSVIRMDRENDLPSELSRREFFVFNKFRQLQKGG 661
QY 148 -HLSPDGQ-----YVPRIM-----FVDPSTVTRADITGRYSNRLYAYE 184
Db 662 FQISNNGQVCFHSFIPKLNDSQYCKPWLRAPEFGYLEPNETV--DIS-----LDYVV 713
QY 185 PADTALLDN 194
Db 714 SKDSVTILNS 723

RESULT 10

US-09-981-353-48
; Sequence 48, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 48
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 1376382CD1
US-09-981-353-48

Query Match 7.3%; Score 78; DB 9; Length 453;
Best Local Similarity 22.5%; Pred. No. 3;
Matches 45; Conservative 34; Mismatches 79; Indels 42; Gaps 9;

QY 25 VRKSRVAMEKIPVSAPFLLLVALSYTLAROTTVKPGAKKDKSRPKLPQLTSLRGWGQDLI 84
Db 187 LRKGFYSQKVTNPNLRIISLNTLYGPNIMTNKTD-----ANQFE 231
QY 85 WTQTYEALYKSTNKNPLMIHH--LDCEPHSQ-----ALKKFAE-----NKEIQKLAEE 133
Db 232 W---LESTLNNQQNKEKYIIIAHVPGVGLPSSQNTAMREYVNEKLLIDIFQKYSVDVIG 288
QY 134 QF---VLLNLVYETTDKHLSPDQGVPRIMFVDPSTVTRADITGRYSN---RLYAYEPA 186
Db 289 QFYGHTHRDSIMVLSDKKGSFVNS-----LFVAPAVTPVKSVLEKOTNPNPGRIFQYDPR 343
QY 187 DTALLDNMKKALKLKTTEL 206
Db 344 DYK-LDDMLQYLYNLTEANL 362

RESULT 11

US-09-815-242-10732
; Sequence 10732, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10732
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10732

Query Match 7.2%; Score 77; DB 10; Length 1049;
Best Local Similarity 25.2%; Pred. No. 12;
Matches 35; Conservative 20; Mismatches 48; Indels 36; Gaps 5;

QY 73 QTLRSGWDQLIWTQTYEALYKSTNKNPLMIHHLDCEPHSQALKKVAENKEIQKLA 132
Db 822 QELNRAY-ESLVTYDDYNDIVQSSTL-----NNQIQLE 855
QY 133 EQFVLLNLVYETT-----DKHLSPDQGVPRIMFVDPSTVTRADITGRYSNRLYAYEPA 187
Db 856 BOIG-----VYETVKGSLJEEDEKEGEDFTIEFYSENWTKLYDIDSTVYDQLLGTYAAN 911
QY 188 TALLDNMMKAL-KLKTKE 205
Db 912 SSDARDEIEKALAKLNKTE 930

RESULT 12

US-09-815-242-13244
; Sequence 13244, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22

Db 143 GEPETDLNPELPAHGLIGAMSGQLSWEASQEEAFRLRSGAQFSP 187

RESULT 15

US-09-849-566-6
 ; Sequence 6, Application US/09849566
 ; Patent No. US20020104109A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bremel, Robert D.
 ; Chan, Anthony W.S.
 ; Burns, Jane C.
 ; TITLE OF INVENTION: Methods For Creating Transgenic Animals
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Medlen & Carroll, LLP
 ; STREET: 220 Montgomery Street, Suite 2200
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States of America
 ; ZIP: 94104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/849,566
 ; FILING DATE: 04-May-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/821,984
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ingolia, Diane E.
 ; REGISTRATION NUMBER: 40,027
 ; REFERENCE/DOCKET NUMBER: WARF-02184
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 705-8410
 ; TELEFAX: (415) 397-8338
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 529 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
 US-09-849-566-6

Query Match 6.7%; Score 71; DB 10; Length 529;
 Best Local Similarity 23.0%; Pred. No. 20;
 Matches 35; Conservative 18; Mismatches 47; Indels 52; Gaps 7;
 QY 19 LTQGRWVRKSRVAMEKIPYSAFLLV-----ALSYTLARDTT-----VKPG 59
 Db 303 LCQNTWDKYSR----KEPLSLDLSVLRAPCKGMAYTVINGTLHSAHAKYIRTWIDYG 358
 QY 60 AKKDKDSRPKLPQTLRSRGWDLIWTQYE-----EALYKSKTSNKLPMII---- 106
 Db 359 EMKEIKGGGRGEYSK-----APELLWSQWFDGPFKIGPNGLLHTGKTFKFPFLYLIGAGI 412
 QY 107 -----HHLDEC-----PHSQALKKKVFAENKEI 128
 Db 413 IDEDLHELDEAAPIDHQPMPDAKSVLPEDDEI 444

Search completed: January 17, 2003, 13:00:53
 Job time : 12 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 12:55:16 ; Search time 15 Seconds
(without alignments)
404.075 Million cell updates/sec

Title: US-09-646-778A-288
Perfect score: 1063
Sequence: 1 RLSCAGTLGSGPHSRRLT.....DTALLDNMKKALKLKTTEL 206

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	899	84.6	175	4	US-08-916-576B-2
2	899	84.6	175	4	US-09-247-155-106
3	609.5	57.3	170	4	US-08-916-576B-8
4	605	56.9	131	4	US-09-247-155-174
5	555	52.2	166	4	US-08-916-576B-6
6	448	42.1	183	4	US-08-916-576B-7
7	222	20.9	172	4	US-08-916-576B-4
8	81.5	7.7	901	2	US-08-884-681-5
9	81.5	7.7	968	3	US-09-258-643-5
10	81.5	7.7	968	3	US-08-560-005-7
11	81.5	7.7	968	4	US-09-418-540-7
12	78	7.3	129	4	US-09-534-638-5
13	78	7.3	453	4	US-09-013-881-5
14	77	7.2	1160	5	PCT-US92-05401-4
15	77	7.2	652	2	US-08-313-185-53
16	77	7.2	652	2	US-08-459-499-17
17	77	7.2	652	3	US-09-082-614A-53
18	74	7.0	264	3	US-09-053-197A-8
19	74	7.0	264	4	US-09-085-761A-8
20	71.5	6.7	352	3	US-08-921-887-52
21	71.5	6.7	382	4	US-09-134-001C-3765
22	71	6.7	529	3	US-08-821-984-6
23	71	6.7	529	3	US-08-821-984-8
24	71	6.7	529	4	US-09-329-749-6
25	71	6.7	529	4	US-09-329-749-8
26	70.5	6.6	481	1	US-08-186-811-2
27	70.5	6.6	481	1	US-08-311-611A-98

28	70.5	6.6	481	1	US-08-372-783-98	Sequence 98, Appl
29	70.5	6.6	481	1	US-08-372-105-98	Sequence 98, Appl
30	70.5	6.6	481	1	US-08-306-473A-98	Sequence 98, Appl
31	70.5	6.6	481	1	US-08-261-660A-4	Sequence 4, Appl
32	70.5	6.6	481	1	US-08-209-762-98	Sequence 98, Appl
33	70.5	6.6	481	1	US-08-473-344-98	Sequence 98, Appl
34	70.5	6.6	481	1	US-08-274-303-6	Sequence 6, Appl
35	70.5	6.6	481	1	US-08-377-391A-2	Sequence 2, Appl
36	70.5	6.6	481	2	US-08-485-445A-98	Sequence 98, Appl
37	70.5	6.6	481	2	US-08-779-400-2	Sequence 2, Appl
38	70.5	6.6	481	2	US-08-955-660-2	Sequence 2, Appl
39	70.5	6.6	481	3	US-09-119-263-98	Sequence 98, Appl
40	70.5	6.6	481	4	US-08-657-162-98	Sequence 98, Appl
41	70.5	6.6	481	4	US-09-224-480-98	Sequence 98, Appl
42	70.5	6.6	481	4	US-09-093-539-98	Sequence 98, Appl
43	70.5	6.6	481	4	US-09-146-620-2	Sequence 2, Appl
44	70.5	6.6	481	4	US-09-395-453-2	Sequence 2, Appl
45	70.5	6.6	481	4	US-09-280-909A-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-916-576B-2
; Sequence 2: Application US/08916576B
; Patent No. 6171816
; GENERAL INFORMATION:
; APPLICANT: YU, GUO-LIANG
; APPLICANT: DILLON, PATRICK J.
; APPLICANT: EBNER, REINHARD
; APPLICANT: ENDRESS, GREGORY A.
; TITLE OF INVENTION: NOVEL HUMAN GROWTH FACTORS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,576B
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,347
; FILING DATE: 23-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0500001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 175 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-916-576B-2

Query Match 84.6%; Score 899; DB 4; Length 175;
Best Local Similarity 100.0%; Pred. No. 1.2e-95;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 32 MEKIPVSFALLVALSYTLARDTVKPKAKKDTKDSRKLPTLSRGWDQLINTQTYEE 91

Db 1 MEKIPVSFLLVLSYTLARDTTVPKAKKDTKDSRPKLPQTLSRGWGDLWTQYEE 60
Qy 92 ALYKSTSNKPLMIHHLEDCPHSQALKKVFVAENKEIQKLAQFVLLNLVYETTDKHLSP 151
Db 61 ALYKSTSNKPLMIHHLEDCPHSQALKKVFVAENKEIQKLAQFVLLNLVYETTDKHLSP 120
Qy 152 DGQYVPRIMFVDPSTVRADITGRYSNRLYAYEPADTALLDNMKKALKLLKTEL 206
Db 121 DGQYVPRIMFVDPSTVRADITGRYSNRLYAYEPADTALLDNMKKALKLLKTEL 175

RESULT 2

US-09-247-155-106
; Sequence 106, Application US/09247155A
; Patent No. 6312922
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET.021A
; CURRENT APPLICATION NUMBER: US/09/247.155A
; EARLIER FILING DATE: 1999-02-09
; EARLIER APPLICATION NUMBER: 60/074.121
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/081.563
; EARLIER FILING DATE: 1998-04-13
; EARLIER APPLICATION NUMBER: 60/096.116
; EARLIER FILING DATE: 1998-08-10
; EARLIER APPLICATION NUMBER: 60/099.273
; EARLIER FILING DATE: 1998-10-04
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 106
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -20..-1
US-09-247-155-106

Query Match 84.6%; Score 899; DB 4; Length 175;
Best Local Similarity 100.0%; Pred. No. 1.2e-95;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 MEKIPVSFLLVLSYTLARDTTVPKAKKDTKDSRPKLPQTLSRGWGDLWTQYEE 91
Db 1 MEKIPVSFLLVLSYTLARDTTVPKAKKDTKDSRPKLPQTLSRGWGDLWTQYEE 60
Qy 92 ALYKSTSNKPLMIHHLEDCPHSQALKKVFVAENKEIQKLAQFVLLNLVYETTDKHLSP 151
Db 61 ALYKSTSNKPLMIHHLEDCPHSQALKKVFVAENKEIQKLAQFVLLNLVYETTDKHLSP 120
Qy 152 DGQYVPRIMFVDPSTVRADITGRYSNRLYAYEPADTALLDNMKKALKLLKTEL 206
Db 121 DGQYVPRIMFVDPSTVRADITGRYSNRLYAYEPADTALLDNMKKALKLLKTEL 175

RESULT 3

US-08-916-576B-8
; Sequence 8, Application US/08916576B
; Patent No. 6171816
; GENERAL INFORMATION:
; APPLICANT: YU, GUO-LIANG
; APPLICANT: DILLON, PATRICK J.
; APPLICANT: EBNER, REINHARD
; APPLICANT: ENDRESS, GREGORY A.
; TITLE OF INVENTION: NOVEL HUMAN GROWTH FACTORS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.

; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,576B
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,347
; FILING DATE: 23-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0500001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-916-576B-8

Query Match 57.3%; Score 609.5; DB 4; Length 170;
Best Local Similarity 71.3%; Pred. No. 2.5e-62;
Matches 122; Conservative 14; Mismatches 26; Indels 9; Gaps 2;

Qy 37 VSAFLLVLSYTLARDTTVPKAKKDTKDSRP-KLPQTLSRGWGDLWTQYEEALYK 95
Db 7 LGACLLVLSXXLA-----XXXKGDXPPIKGPQTLSRGWGDIXVWQTYEGLXK 58
Qy 96 SKTSNKPMLIIHHLEDCPHSQALKKVFVAENKEIQKLAQFVLLNLVYETTDKHLSPDQY 155
Db 59 AXSNKPLMVIHHLEDCPYSQLKKVFVAENKEIQKLAQFVLLNLVYETTDENLSPDQY 118
Qy 156 VPRIMFVDPSTVRADITGRYSNRLYAYEPADTALLDNMKKALKLLKTEL 206
Db 119 VPRIMFVDPSTVRADIXGRYGNRLYAYEPDXPPLIXNMKALKLLKTEL 169

RESULT 4

US-09-247-155-174
; Sequence 174, Application US/09247155A
; Patent No. 6312922
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET.021A
; CURRENT APPLICATION NUMBER: US/09/247.155A
; CURRENT FILING DATE: 1999-02-09
; EARLIER APPLICATION NUMBER: 60/074.121
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/081.563
; EARLIER FILING DATE: 1998-04-13
; EARLIER APPLICATION NUMBER: 60/096.116
; EARLIER FILING DATE: 1998-08-10
; EARLIER APPLICATION NUMBER: 60/099.273
; EARLIER FILING DATE: 1998-10-04
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 174

Db 66 EBLAKARENKPLMVIHLEDCPTYSIALKKAFVADRMAQKLAQEDFIMNLVHPVADEN 125
 QY 149 LSPDQGVPRIMEVDPSLTVRADITGRYSNRLYAYEPADTALLDNMKKALKLTEL 206
 Db 126 QSPDGHYVPRIFIDPSLTVRSDLKGRIGNKMYAYDADDIPELITNMKKAASFLLKTEL 183

RESULT 7

US-08-916-576B-4
 ; Sequence 4, Application US/08916576B
 ; Patent No. 6171816
 ; GENERAL INFORMATION:
 ; APPLICANT: YU, GUO-LIANG
 ; APPLICANT: DILLON, PATRICK J.
 ; APPLICANT: EBNER, REINHARD
 ; APPLICANT: ENDRESS, GREGORY A.
 ; TITLE OF INVENTION: NOVEL HUMAN GROWTH FACTORS
 ; NUMBER OF SEQUENCES: 45
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
 ; STREET: 1100 NEW YORK AVENUE, SUITE 600
 ; CITY: WASHINGTON
 ; STATE: DC
 ; COUNTRY: US
 ; ZIP: 20005-3934

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/916.576B
 FILING DATE:

CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/024,347
 FILING DATE: 23-AUG-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: STEFFE, ERIC K.
 REGISTRATION NUMBER: 36,688
 REFERENCE/DOCKET NUMBER: 1488.0500001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 172 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

US-08-916-576B-4

Query Match 20.9%; Score 222; DB 4; Length 172;
 Best Local Similarity 34.2%; Pred. No. 1.le-17;
 Matches 54; Conservative 26; Mismatches 56; Indels 22; Gaps 3;

QY 66 DSRPKLPQT-----LSRGWDQLIWTQYEEALYKSTNKLPMII 106
 Db 2 ETRPRIGATCGLLGSFLLLYISSDGHNGLGKGFGDHIHW-RTLEDGKRAAASGLPLMVI 60
 QY 107 HHLDECPHSQALKKVFENKEIOKLAEOFVLLNL--VYETTDKHLSPDGOYVPRIMEVDP 164
 Db 61 IHKSWCGACKALPKPKFAESTEISELSNHFVWNLDEDEEPKDEDFSPDGGYIPRILFLDP 120
 QY 165 SLTVRADITGRYSNRLYAYEPADTALLDNMKKALKL 202
 Db 121 SGKVHPPEIINENGNSPKYKYFVSAEQVVGCMKEAQL 158

RESULT 8

US-08-884,681-5
 ; Sequence 5, Application US/08884681

; Patent No. 5955338
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Lal, Preeti
 ; APPLICANT: Corley, Neil C.
 ; APPLICANT: Shah, Purvi
 ; TITLE OF INVENTION: NEW HUMAN PHOSPHATIDYLINOSITOL
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/884,681
 FILING DATE: Filed Herewith
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:

ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0334 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 901 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: GenBank
 CLONE: 1420920

US-08-884-681-5

Query Match 7.7%; Score 81.5; DB 2; Length 901;
 Best Local Similarity 23.7%; Pred. No. 1.8;
 Matches 45; Conservative 32; Mismatches 46; Indels 67; Gaps 13;

QY 55 TVKPGAKDKTSRPRK--LPQTLSRGWDQLIWTQYEEAL-YKS-----KTS-NKPLMII 106
 Db 476 TYKDSKTDWRDSSGKCRVP-----AWCDRIWLWRTNVNQLNYSRSHMELKTSDHKPVSA 530
 QY 107 HHL-----DECPHSQALKKVFENKEIOKLAE-----QFVLLNLVYETTDK-- 147
 Db 531 FHIGVKVDE----RYYKRVFSDSVRIMDRMENDFLPSLSLSRREFVFNKFNQLOK 586
 QY 148 -HLSPDGO-----YVPRIM-----FVDPSTLTVRADITGRYSNRLYAYE 184
 Db 587 FOISNNGQVPCHFSFIPKLNDSQYCKPWLRAEFEGYLEPNEIV--DIS-----LDVYV 638
 QY 185 PADTALLDN 194
 Db 639 SKDSVTILNS 648

RESULT 9

US-09-258-643-5
 ; Sequence 5, Application US/09258643
 ; Patent No. 6277373
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Lal, Preeti
 ; APPLICANT: Corley, Neil C.


```

Db 187 LRKGGYSQKVTNPNLRIISLNLNLYGNNIMTLNKTDP-----ANQFE 231
Qy 85 WTQTYEALYKTSKTNKPLMIHH--LDECPHSQ--ALKKVFAE-----NKEIQKLA 133
Db 232 W---LESTLNNQONKEKVIHIAHPVGYLPSSQNTAMREYNEKLIIDFKYSDVIAG 288
Qy 134 QF---VLLNLVYTTKHLSQDCQYVPRIMFVDSLVVRADITGRYSN-----RLYAYEPA 186
Db 289 QFYGHTRDSIMVLSDKGSPVNS-----LFVAPAVTPVKSLKQTNPNPGRILFOYDPR 343
Qy 187 DTALLDNMKKALKLKLTTEL 206
Db 344 DYK-LLDMLQYLLNLTANL 362

```

RESULT 14

```

PCT-US92-05401-4
; Sequence 4, Application PC/TUS9205401
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ibor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05401
; FILING DATE: 19920626
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1160 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-05401-4

```

```

Query Match 7.3%; Score 78; DB 5; Length 1160;
Best Local Similarity 28.4%; Pred. No. 6.7;
Matches 27; Conservative 17; Mismatches 35; Indels 16; Gaps 4;

Qy 68 RPKLPQTLRSGMGDQLIQTQTYEALYKTSKTNKPLMIHHLDECPHSQALKKVFAPENKE 127
Db 952 RPSFP-NLTSFLGCLQ---ADAEAMYQN-----VDGRVSECPHTYQNRPRPFSREMD 999

Qy 128 IOKLAEQFVL-----LNLVYTTDKHLSPDQYVPR 158
Db 1000 LGLSPQAQVEDSRNNLVLRTSLSLSLTGCRLPK 1034

```

RESULT 15

```

US-08-313-185-53
; Sequence 53, Application US/08313185
; Patent No. 5851763
; GENERAL INFORMATION:
; APPLICANT: Heym, Beate

```

```

; APPLICANT: Cole, Stewart
; APPLICANT: Young, Douglas
; APPLICANT: Zhang, Ying
; APPLICANT: Honore, Nadine
; APPLICANT: Telenti, Amalio
; APPLICANT: Bodmer, Thomas
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
; TITLE OF INVENTION: in Mycobacterium Tuberculosis
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,185
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356.0068-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 652 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-313-185-53

```

```

Query Match 7.2%; Score 77; DB 2; Length 652;
Best Local Similarity 20.7%; Pred. No. 3.7;
Matches 49; Conservative 32; Mismatches 84; Indels 72; Gaps 11;

Qy 12 GPHPSRRLTQGRWVRKSRVAMEKIPYSFALLLVLSYTLARDTTVKPKAKKDTKDSRPLK 71
Db 211 GPNHS-PLSAAAAAIRTRMGMDDETVA---LIAGGHTLGKTHGAGPASHVGVPPAPAP 266
Qy 72 PQTLSRGWGD-----QLIWTQT-----YEEALYKS-----96
Db 267 AAGL--GWASSYSGYGCADAITSGEVVVTQTPTOWNFFENLFYEWVLTSPAGAEAVDG 324
Qy 97 -----KTSNKPMLTIHHLDEC- PHSQALKKVFAPENKEIQKLAEQFVLLNLVYET 144
Db 325 APDIIPDPDPSPKRRKPTMLVTDLLRPEDPEYKISRRLNDPDEFAFARAW-----FKL 378
Qy 145 TDKHLSPD---GOYVPR--IMFVDPPLTVRADI-----TGRYSNRLYAYEPADT 188
Db 379 THRDMGPKRYIGPEVPKEDLIWQDPPQYPTEDIILKAAIAASGLVSELVSAWASAST 435

Search completed: January 17, 2003, 12:56:55
Job time : 17 secs

```

